UNIT I

PHYSICAL LAYER

Data Communications – Networks - Networks models – OSI model – Layers in OSI model TCP / IP protocol suite – Addressing – Guided and Unguided Transmission media Switching: Circuit switched networks – Data gram Networks – Virtual circuit networks Cable networks for Data transmission: Dialup modems – DSL – Cable TV – Cable TV for Data transfer.

1.1.1 Introduction

The concept of Network is not new. In simple terms it means an interconnected set of some objects. For decades we are familiar with the Radio, Television, railway, Highway, Bank and other types of networks. In recent years, the network that is making significant impact in our day-to-day life is the **Computer network**. By computer network we mean an interconnected set of autonomous computers. The term autonomous implies that the computers can function independent of others. However, these computers can exchange information with each other through the communication network system. Computer networks have emerged as a result of the convergence of two technologies of this century- Computer and Communication as shown in Fig. 1.1.1. The consequence of this revolutionary merger is the emergence of a integrated system that transmit all types of data and information. There is no fundamental difference between data communications and data processing and there are no fundamental differences among data, voice and video communications. After a brief historical background in Section 1.1.2, Section 1.1.2 introduces different network categories. A brief overview of the applications of computer networks is presented in Section 1.1.3. Finally an outline of the entire course is given in Section 1.1.4.

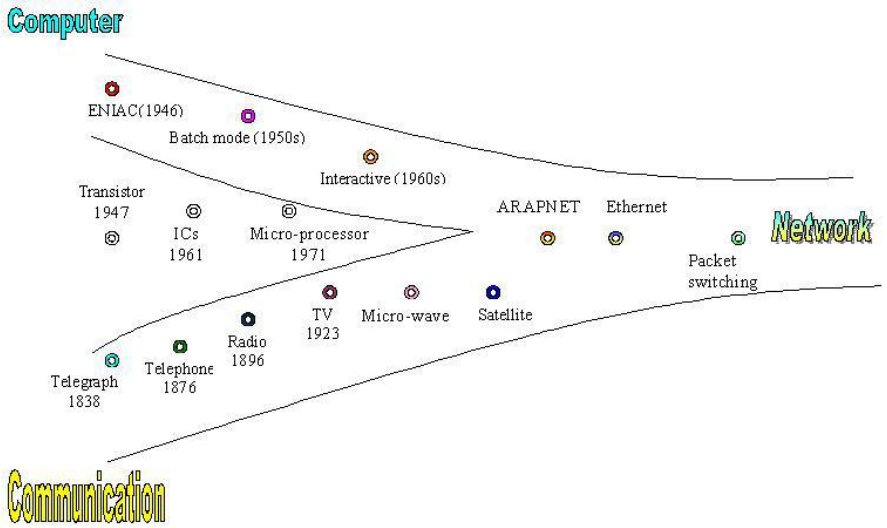


Figure 1.1.1 *Evolution of computer networks*

1.1.2 Historical Background

The history of electronic computers is not very old. It came into existence in the early 1950s and during the first two decades of its existence it remained as a centralized system housed in a single large room. In those days the computers were large in size and were operated by trained personnel. To the users it was a remote and mysterious object having no direct communication with the users. Jobs were submitted in the form of punched cards or paper tape and outputs were collected in the form of computer printouts. The submitted jobs were executed by the computer one after the other, which is referred to as batch mode of data processing. In this scenario, there was long delay between the submission of jobs and receipt of the results.

In the 1960s, computer systems were still centralize, but users provided with direct access through interactive terminals connected by point-to-point low-speed data links with the computer. In this situation, a large number of users, some of them located in remote locations could simultaneously access the centralized computer in time-division multiplexed mode. The users could now get immediate interactive feedback from the computer and correct errors immediately. Following the introduction of on-line terminals and time-sharing operating systems, remote terminals were used to use the central computer.

With the advancement of VLSI technology, and particularly, after the invention of microprocessors in the early 1970s, the computers became smaller in size and less expensive, but with significant increase in processing power. New breed of low-cost computers known as mini and personal computers were introduced. Instead of having a single central computer, an organization could now afford to own a number of computers located in different departments and sections.

Side-by-side, riding on the same VLSI technology the communication technology also advanced leading to the worldwide deployment of telephone network, developed primarily for voice communication. An organization having computers located geographically dispersed locations wanted to have data communications for diverse applications. Communication was required among the machines of the same kind for collaboration, for the use of common software or data or for sharing of some costly resources. This led to the development of computer networks by successful integration and cross-fertilization of communications and geographically dispersed computing facilities. One significant development was the APPANET (Advanced Research Projects Agency Network). Starting with four-node experimental network in 1969, it has subsequently grown into a network several thousand computers spanning half of the globe, from Hawaii to Sweden. Most of the present-day concepts such as packet switching evolved from the ARPANET project. The low bandwidth (3KHz on a voice grade line) telephone network was the only generally available communication system available for this type of network.

The bandwidth was clearly a problem, and in the late 1970s and early 80s another new communication technique known as Local Area Networks (LANs) evolved, which helped computers to communicate at high speed over a small geographical area. In the later years use of optical fiber and satellite communication allowed high-speed data communications over long distances.

1.1.3 Network Technologies

There is no generally accepted taxonomy into which all computer networks fit, but two dimensions stand out as important: **Transmission Technology** and **Scale**. The classifications based on these two basic approaches are considered in this section.

1.1.3.1 Classification Based on Transmission Technology

Computer networks can be broadly categorized into two types based on transmission technologies:

- Broadcast networks
- Point-to-point networks

1.2.3.1.1 Broadcast Networks

Broadcast network have a single communication channel that is shared by all the machines on the network as shown in Figs.1.1.2 and 1.1.3. All the machines on the network receive short messages, called packets in certain contexts, sent by any machine. An address field within the packet specifies the intended recipient. Upon receiving a packet, machine checks the address field. If packet is intended for itself, it processes the packet; if packet is not intended for itself it is simply ignored.

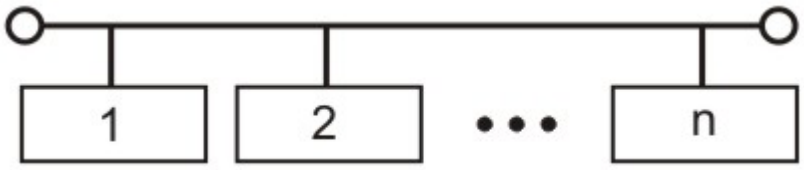


Figure 1.1.2 *Example of a broadcast network based on shared bus*

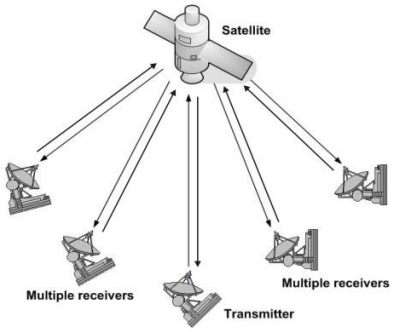


Figure 1.1.3 *Example of a broadcast network based on satellite communication*

This system generally also allows possibility of addressing the packet to all destinations (all nodes on the network). When such a packet is transmitted and received by all the machines on the network. This mode of operation is known as *Broadcast Mode*. Some Broadcast systems also supports transmission to a sub-set of machines, something known as *Multicasting*.

1.2.3.1.2 Point-to-Point Networks

A network based on point-to-point communication is shown in Fig. 1.1.4. The end devices that wish to communicate are called *stations*. The switching devices are called *nodes*. Some Nodes connect to other nodes and some to attached stations. It uses FDM or TDM for node-to-node communication. There may exist multiple paths between a source-destination pair for better network reliability. The switching nodes are not concerned with the contents of data. Their purpose is to provide a switching facility that will move data from node to node until they reach the destination.

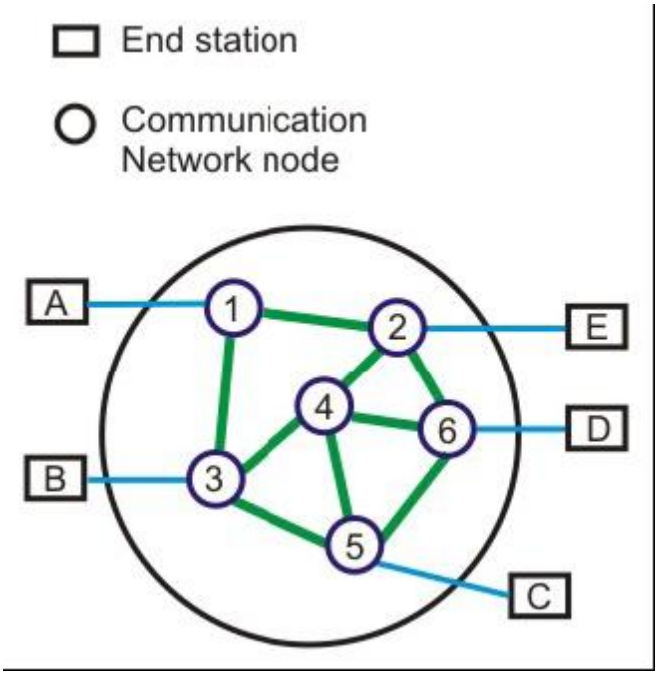


Figure 1.1.4 *Communication network based on point-to-point communication*

As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks normally use are point-to-point communication.

1.1.3.2 Classification based on Scale

Alternative criteria for classifying networks are their scale. They are divided into Local Area (LAN), Metropolitan Area Network (MAN) and Wide Area Networks (WAN).

1.1.3.2.1 Local Area Network (LAN)

LAN is usually privately owned and links the devices in a single office, building or campus of up to few kilometers in size. These are used to share resources (may be hardware or software resources) and to exchange information. LANs are distinguished from other kinds of networks by three categories: their size, transmission technology and topology.

LANs are restricted in size, which means that their worst-case transmission time is bounded and known in advance. Hence this is more reliable as compared to MAN and WAN. Knowing this bound makes it possible to use certain kinds of design that would not otherwise be possible. It also simplifies network management.

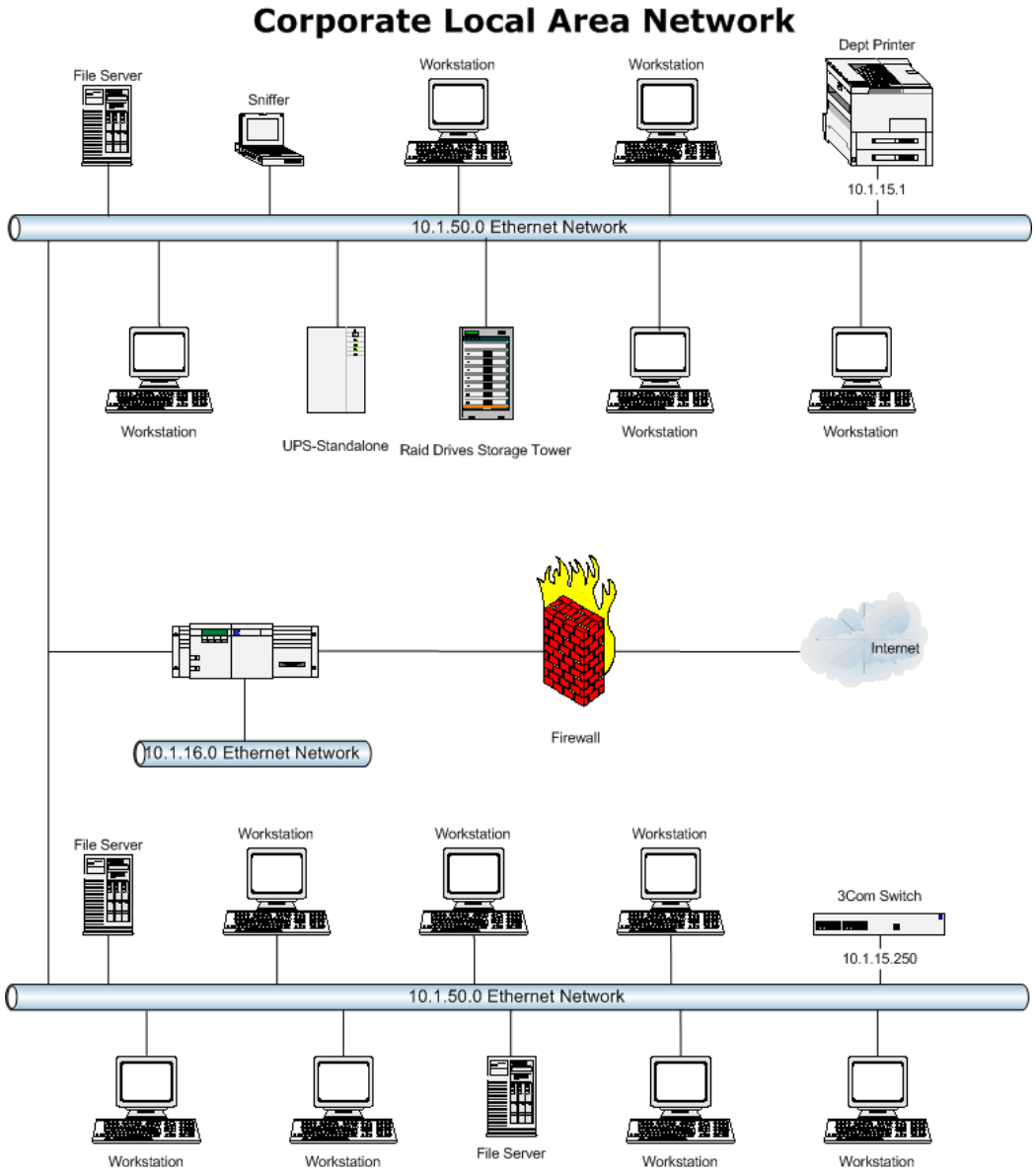


Figure 1.1.5 Local Area Network

LAN typically used transmission technology consisting of single cable to which all machines are connected. Traditional LANs run at speeds of 10 to 100 Mbps (but now much higher speeds can be achieved). The most common LAN topologies are bus, ring and star. A typical LAN is shown in Fig. 1.1.5.

1.1.3.2.2 Metropolitan Area Networks (MAN)

MAN is designed to extend over the entire city. It may be a single network as a cable TV network or it may be means of connecting a number of LANs into a larger network so that resources may be shared as shown in Fig. 1.1.6. For example, a company can use a MAN to connect the LANs in all its offices in a city. MAN is wholly owned and operated by a private company or may be a service provided by a public company.

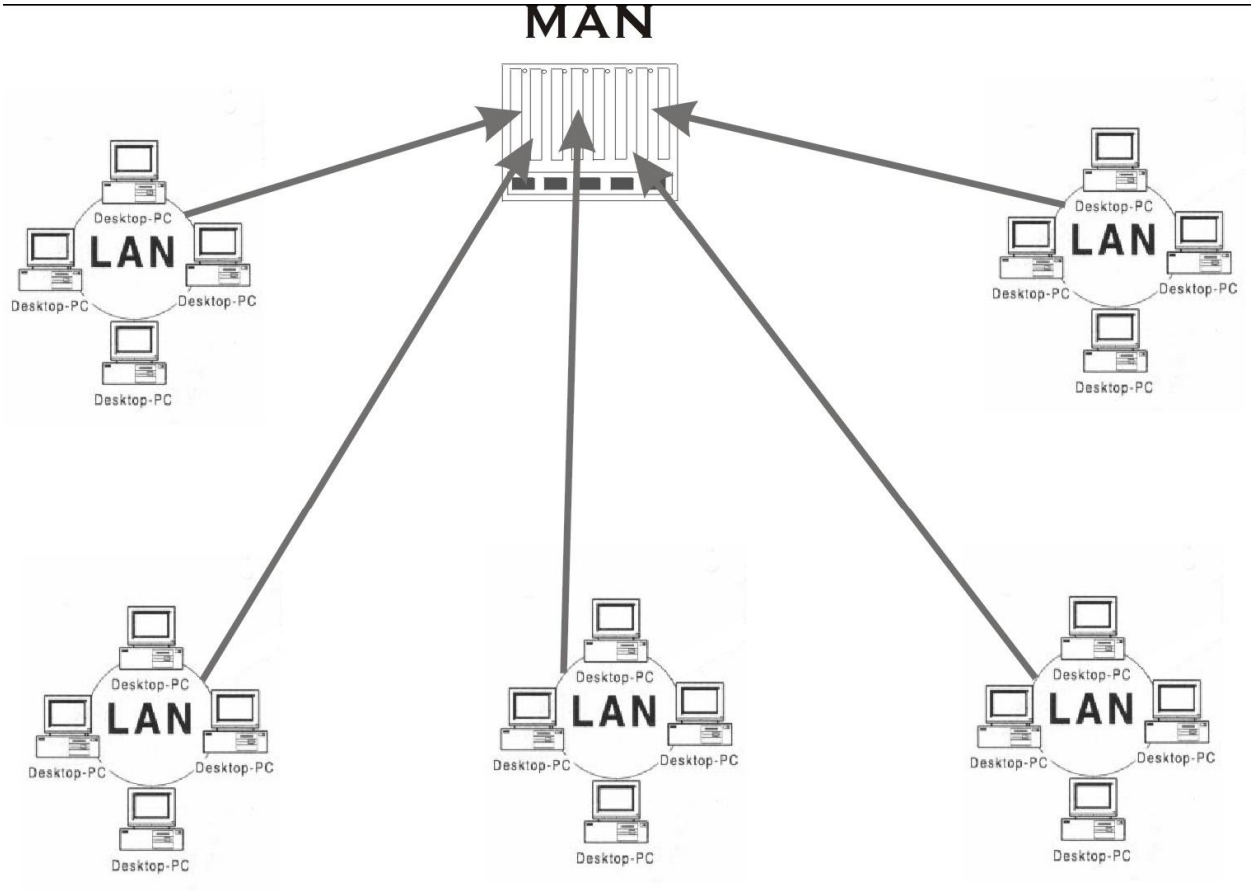


Figure 1.1.6 Metropolitan Area Networks (MAN)

The main reason for distinguishing MANs as a special category is that a standard has been adopted for them. It is **DQDB** (Distributed Queue Dual Bus) or IEEE 802.6.

1.1.3.2.3 Wide Area Network (WAN)

WAN provides long-distance transmission of data, voice, image and information over large geographical areas that may comprise a country, continent or even the whole world. In contrast to LANs, WANs may utilize public, leased or private communication devices, usually in combinations, and can therefore span an unlimited number of miles as shown

in Fig. 1.1.7. A WAN that is wholly owned and used by a single company is often referred to as *enterprise network.*

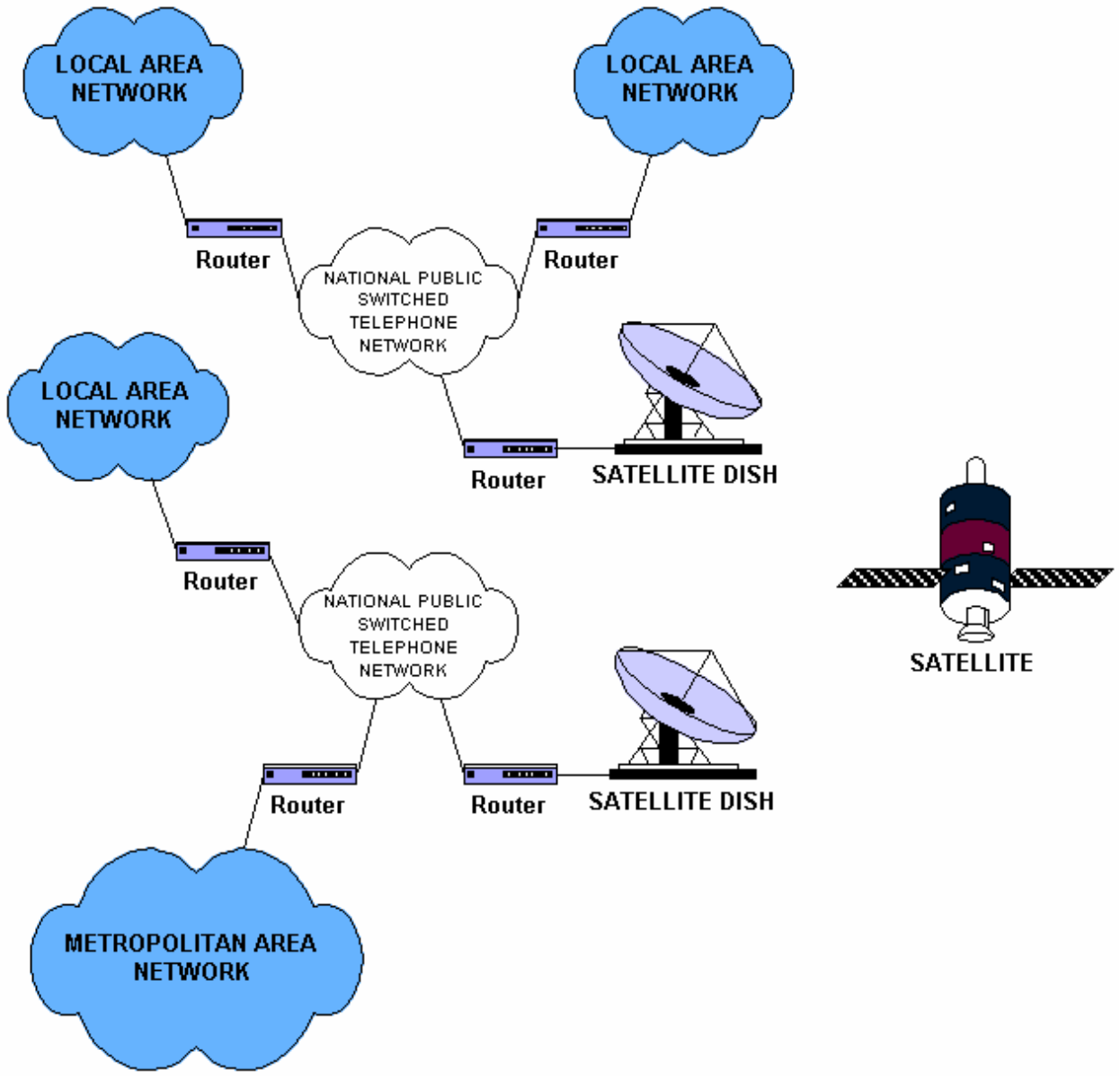


Figure 1.1.7 *Wide Area Network*

1.1.3.2.4 The Internet

Internet is a collection of networks or network of networks. Various networks such as LAN and WAN connected through suitable hardware and software to work in a seamless manner. Schematic diagram of the Internet is shown in Fig. 1.1.8. It allows various applications such as e-mail, file transfer, remote log-in, World Wide Web, Multimedia, etc run across the internet. The basic difference between WAN and Internet is that WAN is owned by a single organization while internet is not so. But with the time the line between WAN and Internet is shrinking, and these terms are sometimes used interchangeably.

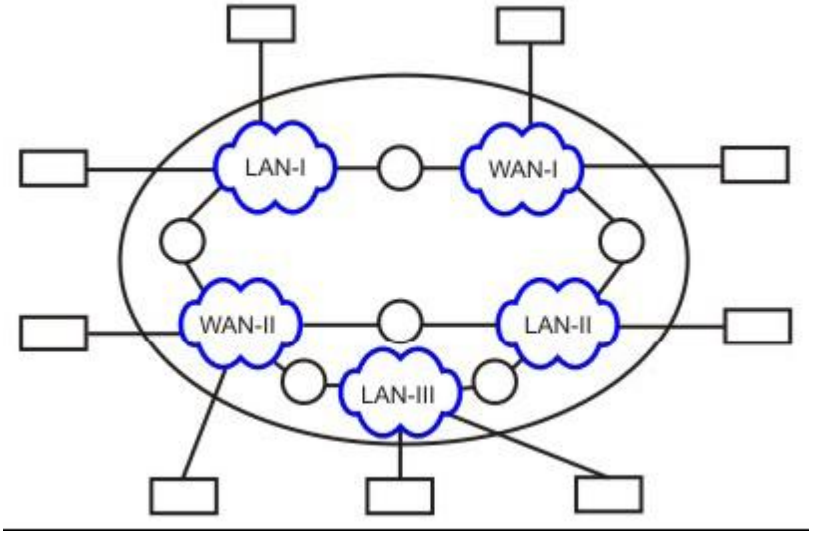


Figure 1.1.8 *Internet – network of networks*

1.1.4 Applications

In a short period of time computer networks have become an indispensable part of business, industry, entertainment as well as a common-man's life. These applications have changed tremendously from time and the motivation for building these networks are all essentially economic and technological.

Initially, computer network was developed for defense purpose, to have a secure communication network that can even withstand a nuclear attack. After a decade or so, companies, in various fields, started using computer networks for keeping track of inventories, monitor productivity, communication between their different branch offices located at different locations. For example, Railways started using computer networks by connecting their nationwide reservation counters to provide the facility of reservation and enquiry from any where across the country.

And now after almost two decades, computer networks have entered a new dimension; they are now an integral part of the society and people. In 1990s, computer network started delivering services to private individuals at home. These services and motivation for using them are quite different. Some of the services are access to remote information, person-person communication, and interactive entertainment. So, some of the applications of computer networks that we can see around us today are as follows:

Marketing and sales: Computer networks are used extensively in both marketing and sales organizations. Marketing professionals use them to collect, exchange, and analyze data related to customer needs and product development cycles. Sales application

includes teleshopping, which uses order-entry computers or telephones connected to order processing network, and online-reservation services for hotels, airlines and so on.

Financial services: Today's financial services are totally depended on computer networks. Application includes credit history searches, foreign exchange and investment services, and electronic fund transfer, which allow user to transfer money without going

into a bank (an automated teller machine is an example of electronic fund transfer, automatic pay-check is another).

Manufacturing: Computer networks are used in many aspects of manufacturing including manufacturing process itself. Two of them that use network to provide essential services are computer-aided design (CAD) and computer-assisted manufacturing (CAM), both of which allow multiple users to work on a project simultaneously.

Directory services: Directory services allow list of files to be stored in central location to speed worldwide search operations.

Information services: A Network information service includes bulletin boards and data banks. A World Wide Web site offering technical specification for a new product is an information service.

Electronic data interchange (EDI): EDI allows business information, including documents such as purchase orders and invoices, to be transferred without using paper.

Electronic mail: probably it's the most widely used computer network application.

Teleconferencing: Teleconferencing allows conference to occur without the participants being in the same place. Applications include simple text conferencing (where participants communicate through their normal keyboards and monitor) and video conferencing where participants can even see as well as talk to other fellow participants. Different types of equipments are used for video conferencing depending on what quality of the motion you want to capture (whether you want just to see the face of other fellow participants or do you want to see the exact facial expression).

Voice over IP: Computer networks are also used to provide voice communication. This kind of voice communication is pretty cheap as compared to the normal telephonic conversation.

Video on demand: Future services provided by the cable television networks may include video on request where a person can request for a particular movie or any clip at anytime he wish to see.

Summary: The main area of applications can be broadly classified into following categories:

Scientific and Technical Computing

– Client Server Model, Distributed Processing

--Parallel Processing, Communication Media

Commercial

– Advertisement, Telemarketing, Teleconferencing

– Worldwide Financial Services

Network for the People (this is the most widely used application nowadays)

– Telemedicine, Distance Education, Access to Remote Information, Person-to-Person Communication, Interactive Entertainment

Layered Network Architecture

Specific Functional Objectives

On Completion of this lesson, the students will be able to:

- State the requirement for layered approach
- Explain the basic concept of layering in the network model
- Define entities protocols in networking context
- Describe ISO's OSI Reference Model
- Explain information flow in OSI references Model.
- Explain functions of the seven layers of OSI Model

1.2.1 Basic concept of layering

Network architectures define the standards and techniques for designing and building communication systems for computers and other devices. In the past, vendors developed their own architectures and required that other vendors conform to this architecture if they wanted to develop compatible hardware and software. There are proprietary network architectures such as IBM's SNA (Systems Network Architecture) and there are open architectures like the OSI (Open Systems Interconnection) model defined by the International Organization for Standardization. The previous strategy, where the computer network is designed with the hardware as the main concern and software is afterthought, no longer works. Network software is now highly *structured.*

To reduce the design complexity, most of the networks are organized as a series of **layers** or **levels**, each one build upon one below it. The basic idea of a layered architecture is *to divide the design into small pieces*. Each layer adds to the services provided by the lower layers in such a manner that the highest layer is provided a full set of services to manage communications and run the applications. The benefits of the layered models are modularity and clear interfaces, i.e. open architecture and comparability between the different providers' components.

A basic principle is to ensure independence of layers by defining services provided by each layer to the next higher layer without defining how the services are to be performed. This permits changes in a layer without affecting other layers. Prior to the use of layered protocol architectures, simple changes such as adding one terminal type to the list of those supported by an architecture often required changes to essentially all communications software at a site. The number of layers, functions and contents of each layer differ from network to network. However in all networks, the purpose of each layer is to offer certain services to higher layers, shielding those layers from the details of how the services are actually implemented.

The basic elements of a layered model are services, protocols and interfaces. A *service* is a set of actions that a layer offers to another (higher) layer. *Protocol* is a set of rules that a layer uses to exchange information with a peer entity. These rules concern both the contents and the order of the messages used. Between the layers service interfaces are defined. The messages from one layer to another are sent through those interfaces.

In an n-layer architecture, layer n on one machine carries on conversation with the layer n on other machine. The rules and conventions used in this conversation are collectively known as the *layer-n protocol*. Basically, a protocol is an agreement between the communicating parties on how communication is to proceed. Violating the protocol will make communication more difficult, if not impossible. A five-layer architecture is shown in Fig. 1.2.1**,** the entities comprising the corresponding layers on different machines are called *peers*. In other words, it is the peers that communicate using protocols. In reality, no data is transferred from layer n on one machine to layer n of another machine. Instead, each layer passes data and control information to the layer immediately below it, until the lowest layer is reached. Below layer-1 is the physical layer through which actual communication occurs. The peer process abstraction is crucial to all network design. Using it, the un-manageable tasks of designing the complete network can be broken into several smaller, manageable, design problems, namely design of individual layers.

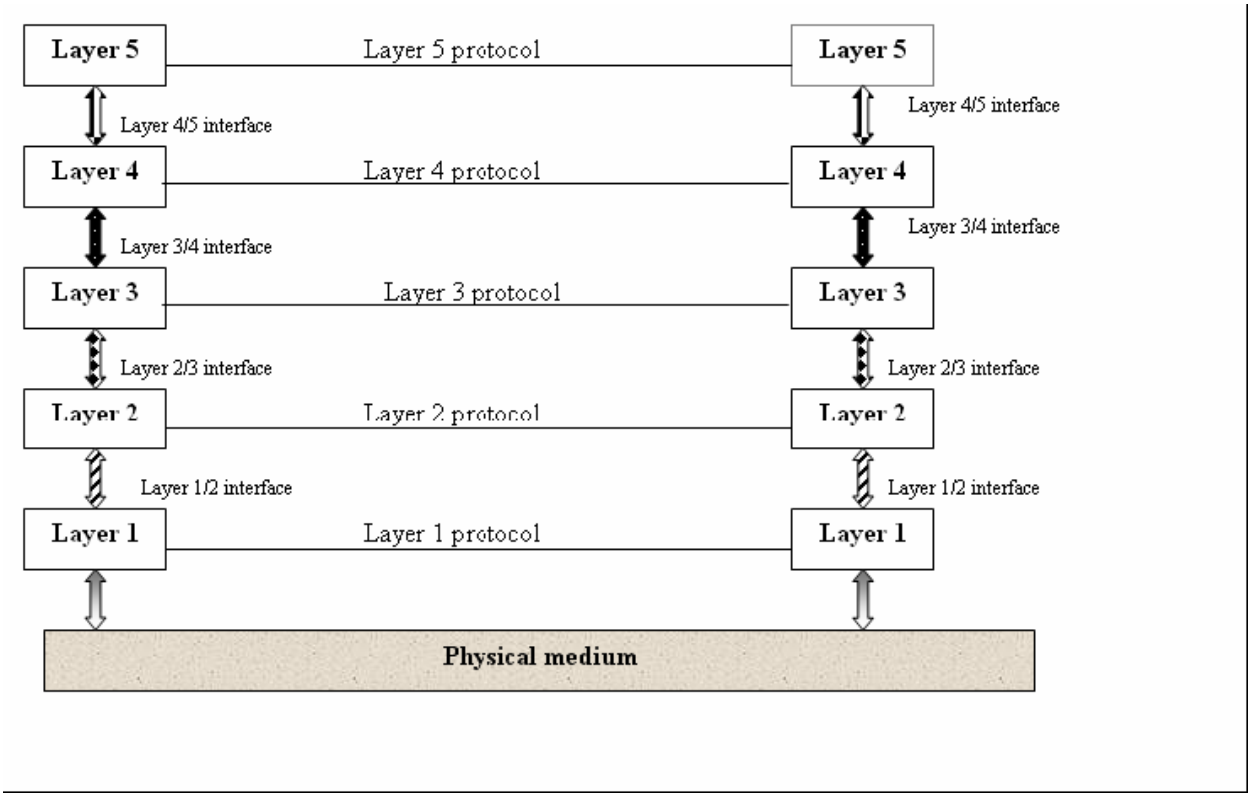


Figure 1.2.1 *Basic five layer architecture*

Between each pair of adjacent layers there is an **interface**. The *interface* defines which primitives operations and services the lower layer offers to the upper layer adjacent to it. When network designer decides how many layers to include in the network and what each layer should do, one of the main considerations is defining clean interfaces between adjacent layers. Doing so, in turns requires that each layer should perform well-defined functions. In addition to minimize the amount of information passed between layers,

clean-cut interface also makes it simpler to replace the implementation of one layer with a completely different implementation, because all what is required of new implementation is that it offers same set of services to its upstairs neighbor as the old implementation (that is what a layer provides and how to use that service from it is more important than knowing how exactly it implements it).

A set of layers and protocols is known as **network architecture**. The specification of architecture must contain enough information to allow an implementation to write the program or build the hardware for each layer so that it will correctly obey the appropriate protocol. Neither the details of implementation nor the specification of interface is a part of network architecture because these are hidden away inside machines and not visible from outside. It is not even necessary that the interface on all machines in a network be same, provided that each machine can correctly use all protocols. A list of protocols used by a certain system, one protocol per layer, is called **protocol stack**.

Summary: Why Layered architecture?

1. To make the design process easy by breaking unmanageable tasks into several smaller and manageable tasks (by divide-and-conquer approach).

2. Modularity and clear interfaces, so as to provide comparability between the different providers' components.

3. Ensure independence of layers, so that implementation of each layer can be changed or modified without affecting other layers.

4. Each layer can be analyzed and tested independently of all other layers.

1.2.2 Open System Interconnection Reference Model

The Open System Interconnection (OSI) reference model describes how information from a software application in one computer moves through a network medium to a software application in another computer. The OSI reference model is a conceptual model composed of seven layers, each specifying particular network functions. The model was developed by the International Organization for Standardization (ISO) in 1984, and it is now considered the primary architectural model for inter-computer communications. The OSI model divides the tasks involved with moving information between networked computers into seven smaller, more manageable task groups. A task or group of tasks is then assigned to each of the seven OSI layers. Each layer is reasonably self-contained so that the tasks assigned to each layer can be implemented independently. This enables the solutions offered by one layer to be updated without adversely affecting the other layers.

The OSI Reference Model includes seven layers:

7. *Application Layer*: Provides Applications with access to network services.

6. *Presentation Layer*: Determines the format used to exchange data among networked computers.

5. *Session Layer*: Allows two applications to establish, use and disconnect a connection between them called a session. Provides for name recognition and additional functions like security, which are needed to allow applications to communicate over the network.

4. *Transport Layer*: Ensures that data is delivered error free, in sequence and with no loss, duplications or corruption. This layer also repackages data by assembling long messages into lots of smaller messages for sending, and repackaging the smaller messages into the original larger message at the receiving end.

3. *Network Layer*: This is responsible for addressing messages and data so they are sent to the correct destination, and for translating logical addresses and names (like a machine name FLAME) into physical addresses. This layer is also responsible for finding a path through the network to the destination computer.

2. *Data-Link Layer*: This layer takes the data frames or messages from the Network Layer and provides for their actual transmission. At the receiving computer, this layer receives the incoming data and sends it to the network layer for handling. The Data-Link Layer also provides error-free delivery of data between the two computers by using the physical layer. It does this by packaging the data from the Network Layer into a frame, which includes error detection information. At the receiving computer, the Data-Link Layer reads the incoming frame, and generates its own error detection information based on the received frames data. After receiving the entire frame, it then compares its error detection value with that of the incoming frames, and if they match, the frame has been received correctly.

1. *Physical Layer*: Controls the transmission of the actual data onto the network cable. It defines the electrical signals, line states and encoding of the data and the connector types used. An example is 10BaseT.

1.2.2.1Characteristics of the OSI Layers

The seven layers of the OSI reference model can be divided into two categories: upper layers and lower layers as shown in Fig. 1.2.2.

The upper layers of the OSI model deal with application issues and generally are implemented only in software. The highest layer, the application layer, is closest to the end user. Both users and application layer processes interact with software applications that contain a communications component. The term upper layer is sometimes used to refer to any layer above another layer in the OSI model.

The lower layers of the OSI model handle data transport issues. The physical layer and the data link layer are implemented in hardware and software. The lowest layer, the physical layer, is closest to the physical network medium (the network cabling, for example) and is responsible for actually placing information on the medium .

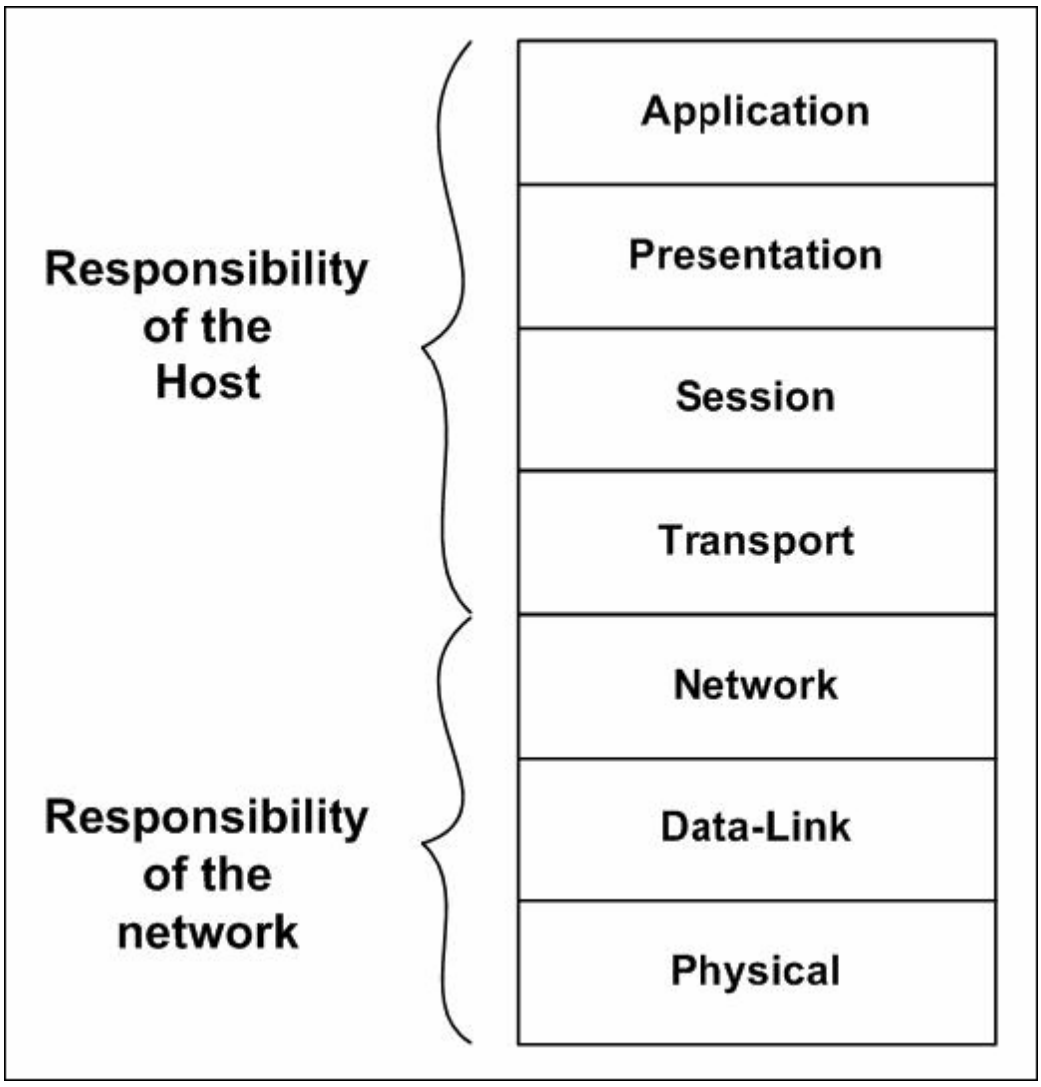


Figure 1.2.2 *Two sets of layers make up the OSI layers*

1.2.2.2 Protocols

The OSI model provides a conceptual framework for communication between computers, but the model itself is not a method of communication. Actual communication is made possible by using communication protocols. In the context of data networking, a **protocol** *is a formal set of rules and conventions that governs how computers exchange information over a network medium*. A protocol implements the functions of one or more of the OSI layers.

A wide variety of communication protocols exist. Some of these protocols include LAN protocols, WAN protocols, network protocols, and routing protocols. LAN protocols operate at the physical and data link layers of the OSI model and define communication over various LAN media. WAN protocols operate at the lowest three layers of the OSI model and define communication over the various wide-area media. Routing protocols

are network layer protocols that are responsible for exchanging information between routers so that the routers can select the proper path for network traffic. Finally, network protocols are the various upper-layer protocols that exist in a given protocol suite. Many protocols rely on others for operation. For example, many routing protocols use network protocols to exchange information between routers. This concept of building upon the layers already in existence is the foundation of the OSI model.

1.2.2.3 OSI Model and Communication between Systems

Information being transferred from a software application in one computer system to a software application in another must pass through the OSI layers. For example, if a software application in System A has information to transmit to a software application in System B, the application program in System A will pass its information to the application layer (Layer 7) of System A. The application layer then passes the information to the presentation layer (Layer 6), which relays the data to the session layer (Layer 5), and so on down to the physical layer (Layer 1). At the physical layer, the information is placed on the physical network medium and is sent across the medium to System B. The physical layer of System B removes the information from the physical medium, and then its physical layer passes the information up to the data link layer (Layer 2), which passes it to the network layer (Layer 3), and so on, until it reaches the application layer (Layer 7) of System B. Finally, the application layer of System B passes the information to the recipient application program to complete the communication process.

1.2.2.4 Interaction between OSI Model Layers

A given layer in the OSI model generally communicates with three other OSI layers: the layer directly above it, the layer directly below it, and its peer layer in other networked computer systems. The data link layer in System A, for example, communicates with the network layer of System A, the physical layer of System A, and the data link layer in System B. Figure1.2.3 illustrates this example.

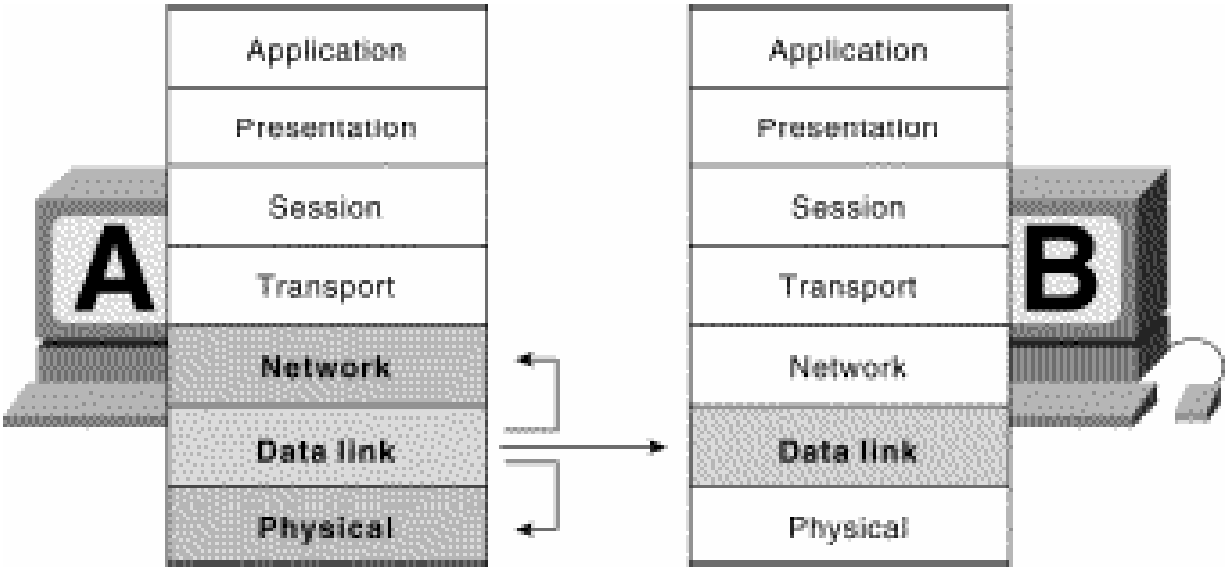


Figure 1.2.3 *OSI Model Layers Communicate with Other Layers*

1.2.3 Services and service access points

One OSI layer communicates with another layer to make use of the services provided by the second layer. The services provided by adjacent layers help a given OSI layer communicate with its peer layer in other computer systems. Three basic elements are involved in layer services: the service user, the service provider, and the service access point (SAP).

In this context, the service user is the OSI layer that requests services from an adjacent OSI layer. The service provider is the OSI layer that provides services to service users. OSI layers can provide services to multiple service users. The SAP is a conceptual location at which one OSI layer can request the services of another OSI layer.

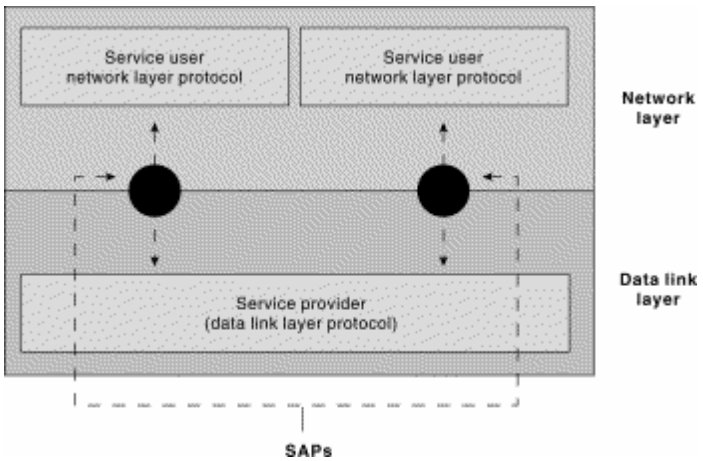


Figure 1.2.4 *Service Users, Providers, and SAPs interact at the Network and Data Link Layers*

1.2.3.1 OSI Model Layers and Information Exchange

The seven OSI layers use various forms of control information to communicate with their peer layers in other computer systems. This control information consists of specific requests and instructions that are exchanged between peer OSI layers.

Control information typically takes one of two forms: headers and trailers. Headers are prepended to data that has been passed down from upper layers. Trailers are appended to data that has been passed down from upper layers. An OSI layer is not required to attach a header or a trailer to data from upper layers.

Headers, trailers, and data are relative concepts, depending on the layer that analyzes the information unit. At the network layer, for example, an information unit consists of a Layer 3 header and data. At the data link layer, however, all the information passed down by the network layer (the Layer 3 header and the data) is treated as data.

In other words, the data portion of an information unit at a given OSI layer potentially can contain headers, trailers, and data from all the higher layers. This is known as

encapsulation. Figure 1-6 shows how the header and data from one layer are encapsulated into the header of the next lowest layer.

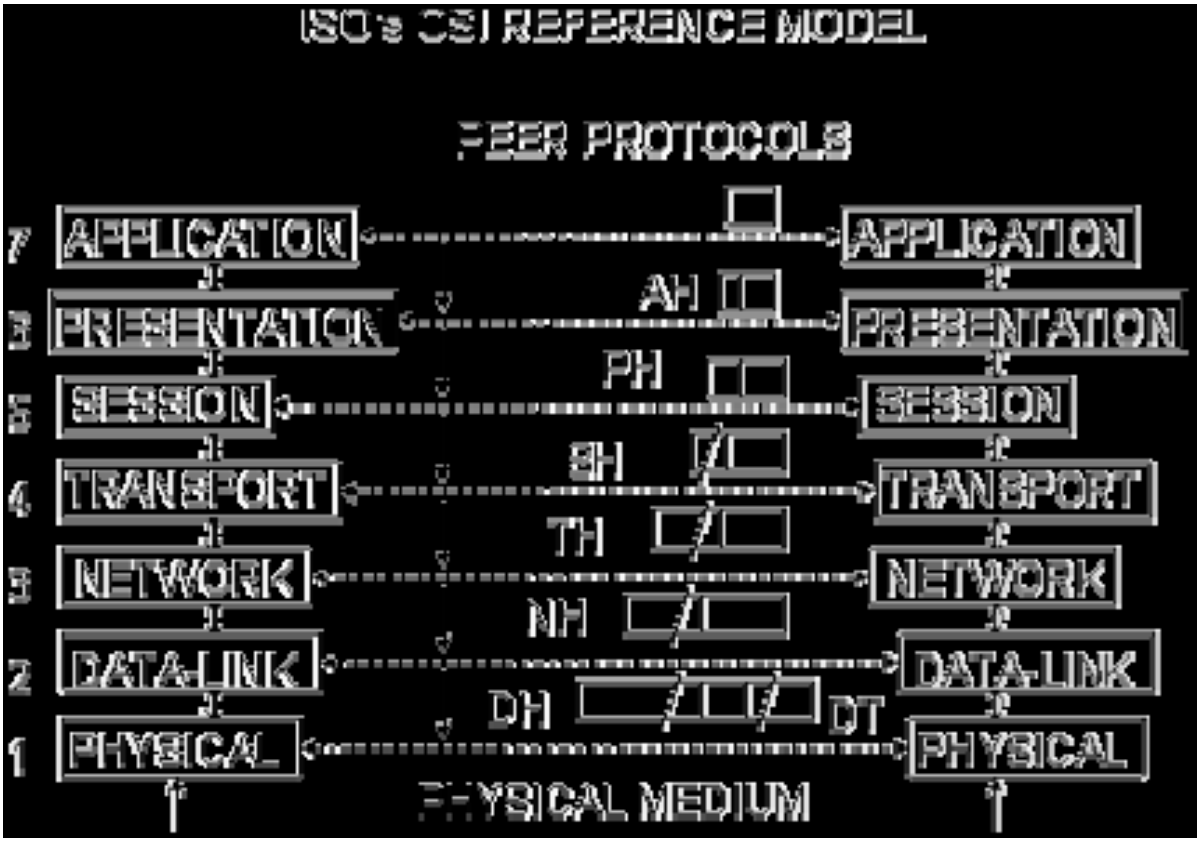


Figure 1.2.6 *Headers and Data can be encapsulated during Information exchange*

1.2.3.2 Information Exchange Process

The information exchange process occurs between peer OSI layers. Each layer in the source system adds control information to data, and each layer in the destination system analyzes and removes the control information from that data.

If system A has data from software application to send to System B, the data is passed to the application layer. The application layer in System A then communicates any control information required by the application layer in System B by pre-pending a header to the data. The resulting information unit (a header and the data) is passed to the presentation layer, which pre-pends its own header containing control information intended for the presentation layer in System B. The information unit grows in size as each layer pre-pends its own header (and, in some cases, a trailer) that contains control information to be used by its peer layer in System B. At the physical layer, the entire information unit is placed onto the network medium.

The physical layer in System B receives the information unit and passes it to the data link layer. The data link layer in System B then reads the control information contained in the

header pre-pended by the data link layer in System A. The header is then removed, and the remainder of the information unit is passed to the network layer. Each layer performs the same actions: The layer reads the header from its peer layer, strips it off, and passes the remaining information unit to the next highest layer. After the application layer performs these actions, the data is passed to the recipient software application in System B, in exactly the form in which it was transmitted by the application in System A.

1.2.4 Functions of the OSI Layers

Functions of different layers of the OSI model are presented in this section.

1.2.4.1 Physical Layer

The physical layer is concerned with transmission of raw bits over a communication channel. It specifies the mechanical, electrical and procedural network interface specifications and the physical transmission of bit streams over a transmission medium connecting two pieces of communication equipment. In simple terns, the physical layer decides the following:

- Number of pins and functions of each pin of the network connector (Mechanical)
- Signal Level, Data rate (Electrical)
- Whether simultaneous transmission in both directions
- Establishing and breaking of connection
- Deals with physical transmission

There exist a variety of physical layer protocols such as RS-232C, Rs-449 standards developed by Electronics Industries Association (EIA).

1.2.4.2 Data Link Layer

The goal of the data link layer is to provide reliable, efficient communication between adjacent machines connected by a single communication channel. Specifically:

1. Group the physical layer bit stream into units called frames. Note that frames are nothing more than ``packets" or ``messages". By convention, we shall use the term ``frames" when discussing DLL packets.

2. Sender calculates the checksum and sends checksum together with data. The checksum allows the receiver to determine when a frame has been damaged in transit or received correctly.

3. Receiver recomputes the checksum and compares it with the received value. If they differ, an error has occurred and the frame is discarded.

4. Error control protocol returns a positive or negative acknowledgment to the sender. A positive acknowledgment indicates the frame was received without errors, while a negative acknowledgment indicates the opposite.

5. Flow control prevents a fast sender from overwhelming a slower receiver. For example, a supercomputer can easily generate data faster than a PC can consume it.

6. In general, data link layer provides service to the network layer. The network layer wants to be able to send packets to its neighbors without worrying about the details of getting it there in one piece.

1.2.4.2.1 Design Issues Below are the some of the important design issues of the data link layer:

a). Reliable Delivery:

Frames are delivered to the receiver reliably and in the same order as generated by the sender. Connection state keeps track of sending order and which frames require retransmission. For example, receiver state includes which frames have been received, which ones have not, etc.

b). Best Effort: The receiver does not return acknowledgments to the sender, so the sender has no way of knowing if a frame has been successfully delivered.

When would such a service be appropriate?

1. When higher layers can recover from errors with little loss in performance. That is, when errors are so infrequent that there is little to be gained by the data link layer performing the recovery. It is just as easy to have higher layers deal with occasional loss of packet.

2. For real-time applications requiring ``better never than late" semantics. Old data may be worse than no data.

c). Acknowledged Delivery

The receiver returns an acknowledgment frame to the sender indicating that a data frame was properly received. This sits somewhere between the other two in that the sender keeps connection state, but may not necessarily retransmit unacknowledged frames. Likewise, the receiver may hand over received packets to higher layer in the order in which they arrive, regardless of the original sending order. Typically, each frame is assigned a unique sequence number, which the receiver returns in an acknowledgment frame to indicate which frame the ACK refers to. The sender must retransmit unacknowledged (e.g., lost or damaged) frames.

d). Framing

The DLL translates the physical layer's raw bit stream into discrete units (messages) called *frames*. How can the receiver detect frame boundaries? Various techniques are used for this: Length Count, Bit Stuffing, and Character stuffing.

e). Error Control

Error control is concerned with insuring that all frames are eventually delivered (possibly in order) to a destination. To achieve this, three items are required: Acknowledgements, Timers, and Sequence Numbers.

f). Flow Control

Flow control deals with throttling the speed of the sender to match that of the receiver. Usually, this is a dynamic process, as the receiving speed depends on such changing factors as the load, and availability of buffer space.

1.2.4.2.2 Link Management In some cases, the data link layer service must be ``opened'' before use:

• The data link layer uses open operations for allocating buffer space, control blocks, agreeing on the maximum message size, etc.

• Synchronize and initialize send and receive sequence numbers with its peer at the other end of the communications channel.

1.2.4.2.3 Error Detection and Correction

In data communication, error may occur because of various reasons including attenuation, noise. Moreover, error usually occurs as bursts rather than independent, single bit errors. For example, a burst of lightning will affect a set of bits for a short time after the lightning strike. Detecting and correcting errors requires redundancy (i.e., sending additional information along with the data).

There are two types of attacks against errors:

• Error Detecting Codes: Include enough redundancy bits to detect errors and use ACKs and retransmissions to recover from the errors. Example: parity encoding.

• Error Correcting Codes: Include enough redundancy to detect and correct errors. Examples: CRC checksum, MD5.

1.2.4.3 Network Layer

The basic purpose of the network layer is to provide an end-to-end communication capability in contrast to machine-to-machine communication provided by the data link layer. This end-to-end is performed using two basic approaches known as connection-oriented or connectionless network-layer services.

1.2.4.3.1 Four issues:

1. Interface between the host and the network (the network layer is typically the boundary between the host and subnet)

2. Routing

3. Congestion and deadlock

4. Internetworking (A path may traverse different network technologies (e.g., Ethernet, point-to-point links, etc.)

1.2.4.3.2 Network Layer Interface

There are two basic approaches used for sending packets, which is a group of bits that includes data plus source and destination addresses, from node to node called *virtual circuit* and *datagram* methods. These are also referred to as *connection-oriented* and *connectionless* network-layer services. In virtual circuit approach, a *route*, which consists

of logical connection, is first established between two users. During this establishment phase, the two users not only agree to set up a connection between them but also decide upon the quality of service to be associated with the connection. The well-known virtual-circuit protocol is the ISO and CCITT *X.25* specification. The datagram is a self-contained message unit, which contains sufficient information for routing from the source node to the destination node without dependence on previous message interchanges between them. In contrast to the virtual-circuit method, where a fixed path is explicitly set up before message transmission, sequentially transmitted messages can follow completely different paths. The datagram method is analogous to the postal system and the virtual-circuit method is analogous to the telephone system.

1.2.4.3.3 Overview of Other Network Layer Issues:

The network layer is responsible for routing packets from the source to destination. The *routing algorithm* is the piece of software that decides where a packet goes next (e.g., which output line, or which node on a broadcast channel).

For connectionless networks, the routing decision is made for each datagram. For connection-oriented networks, the decision is made once, at circuit setup time.

1.2.4.3.4 Routing Issues:

The routing algorithm must deal with the following issues:

• Correctness and simplicity: networks are never taken down; individual parts (e.g., links, routers) may fail, but the whole network should not.

• Stability: if a link or router fails, how much time elapses before the remaining routers recognize the topology change? (Some never do.)

• Fairness and optimality: an inherently intractable problem. Definition of optimality usually doesn't consider fairness. Do we want to maximize channel usage? Minimize average delay?

When we look at routing in detail, we'll consider both adaptive--those that take current traffic and topology into consideration--and non-adaptive algorithms.

1.2.4.3.4 Congestion The network layer also must deal with congestion:

• When more packets enter an area than can be processed, delays increase and performance decreases. If the situation continues, the subnet may have no alternative but to discard packets.

• If the delay increases, the sender may (incorrectly) retransmit, making a bad situation even worse.

• Overall, performance degrades because the network is using (wasting) resources processing packets that eventually get discarded.

1.2.4.3.5 Internetworking Finally, when we consider internetworking -- connecting different network technologies together -- one finds the same problems, only worse:

• Packets may travel through many different networks

• Each network may have a different frame format

• Some networks may be connectionless, other connection oriented

1.2.4.3.6 Routing

Routing is concerned with the question: Which line should router J use when forwarding a packet to router K?

There are two types of algorithms:

• **Adaptive algorithms** use such dynamic information as current topology, load, delay, etc. to select routes.

• In **non-adaptive algorithms**, routes never change once initial routes have been selected. Also called static routing.

Obviously, adaptive algorithms are more interesting, as non-adaptive algorithms don't even make an attempt to handle failed links.

1.2.4.4 Transport Layer

The transport level provides end-to-end communication between processes executing on different machines. Although the services provided by a transport protocol are similar to those provided by a data link layer protocol, there are several important differences between the transport and lower layers:

1. User Oriented. Application programmers interact directly with the transport layer, and from the programmers perspective, the transport layer is the ``network''. Thus, the transport layer should be oriented more towards user services than simply reflect what the underlying layers happen to provide. (Similar to the beautification principle in operating systems.)

2. Negotiation of Quality and Type of Services. The user and transport protocol may need to negotiate as to the quality or type of service to be provided. Examples? A user may want to negotiate such options as: throughput, delay, protection, priority, reliability, etc.

3. Guarantee Service. The transport layer may have to overcome service deficiencies of the lower layers (e.g. providing reliable service over an unreliable network layer).

4. Addressing becomes a significant issue. That is, now the user must deal with it; before it was buried in lower levels.

Two solutions:

• Use well-known addresses that rarely if ever change, allowing programs to ``wire in'' addresses. For what types of service does this work? While this works for services that are well established (e.g., mail, or telnet), it doesn't allow a user to easily experiment with new services.

• Use a name server. Servers register services with the name server, which clients contact to find the transport address of a given service.

In both cases, we need a mechanism for mapping high-level service names into low-level encoding that can be used within packet headers of the network protocols. In its general

form, the problem is quite complex. One simplification is to break the problem into two parts: have transport addresses be a combination of machine address and local process on that machine.

***5. Storage capacity* of the subnet.** Assumptions valid at the data link layer do not necessarily hold at the transport Layer. Specifically, the subnet may buffer messages for a potentially long time, and an ``old'' packet may arrive at a destination at unexpected times.

6. We need a dynamic flow control mechanism. The data link layer solution of reallocating buffers is inappropriate because a machine may have hundreds of connections sharing a single physical link. In addition, appropriate settings for the flow control parameters depend on the communicating end points (e.g., Cray supercomputers vs. PCs), not on the protocol used.

Don't send data unless there is room. Also, the network layer/data link layer solution of simply not acknowledging frames for which the receiver has no space is unacceptable. Why? In the data link case, the line is not being used for anything else; thus retransmissions are inexpensive. At the transport level, end-to-end retransmissions are needed, which wastes resources by sending the same packet over the same links multiple times. If the receiver has no buffer space, the sender should be prevented from sending data.

7. *Deal with congestion control*. In connectionless Internets, transport protocols must exercise congestion control. When the network becomes congested, they must reduce rate at which they insert packets into the subnet, because the subnet has no way to prevent itself from becoming overloaded.

8. *Connection establishment.* Transport level protocols go through three phases: establishing, using, and terminating a connection. For data gram-oriented protocols, opening a connection simply allocates and initializes data structures in the operating system kernel.

Connection oriented protocols often exchanges messages that negotiate options with the remote peer at the time a connection are opened. Establishing a connection may be tricky because of the possibility of old or duplicate packets.

Finally, although not as difficult as establishing a connection, terminating a connection presents subtleties too. For instance, both ends of the connection must be sure that all the data in their queues have been delivered to the remote application.

1.2.4.5 Session Layer

This layer allows users on different machines to establish session between them. A session allows ordinary data transport but it also provides enhanced services useful in some applications. A session may be used to allow a user to log into a remote time-

sharing machine or to transfer a file between two machines. Some of the session related services are:

1. This layer manages *Dialogue Control*. Session can allow traffic to go in both direction at the same time, or in only one direction at one time.

2. *Token management.* For some protocols, it is required that both sides don't attempt same operation at the same time. To manage these activities, the session layer provides tokens that can be exchanged. Only one side that is holding token can perform the critical operation. This concept can be seen as entering into a critical section in operating system using semaphores.

3. *Synchronization*. Consider the problem that might occur when trying to transfer a 4-hour file transfer with a 2-hour mean time between crashes. After each transfer was aborted, the whole transfer has to start again and again would probably fail. To Eliminate this problem, Session layer provides a way to insert checkpoints into data streams, so that after a crash, only the data transferred after the last checkpoint have to be repeated.

1.2.4.6 Presentation Layer

This layer is concerned with Syntax and Semantics of the information transmitted, unlike other layers, which are interested in moving data reliably from one machine to other. Few of the services that Presentation layer provides are:

1. Encoding data in a standard agreed upon way.

2. It manages the abstract data structures and converts from representation used inside computer to network standard representation and back.

1.2.4.7 Application Layer

The application layer consists of what most users think of as programs. The application does the actual work at hand. Although each application is different, some applications are so useful that they have become standardized. The Internet has defined standards for:

• File transfer (FTP): Connect to a remote machine and send or fetch an arbitrary file. FTP deals with authentication, listing a directory contents, ASCII or binary files, etc.

• Remote login (telnet): A remote terminal protocol that allows a user at one site to establish a TCP connection to another site, and then pass keystrokes from the local host to the remote host.

• Mail (SMTP): Allow a mail delivery agent on a local machine to connect to a mail delivery agent on a remote machine and deliver mail.

• News (NNTP): Allows communication between a news server and a news client.

• Web (HTTP): Base protocol for communication on the World Wide Web.

Review questions

Q-1. Why it is necessary to have layering in a network?

Ans: A computer network is a very complex system. It becomes very difficult to implement as a single entity. The layered approach divides a very complex task into small pieces each of which is independent of others and it allow a structured approach in implementing a network. The basic idea of a layered architecture is *to divide the design into small pieces*. Each layer adds to the services provided by the lower layers in such a

manner that the highest layer is provided a full set of services to manage communications and run the applications.

Q-2. What are the key benefits of layered network?

Ans: Main benefits of layered network are given below:

i) Complex systems can be broken down into understandable subsystems.
ii) Any facility implemented in one layer can be made visible to all other layers.
iii)Services offered at a particular level may share the services of lower level.
iv)Each layer may be analyzed and tested independently.
v) Layers can be simplified, extended or deleted at any time.
vi)Increase the interoperability and compatibility of various components build by different vendors.

Q-3. What do you mean by OSI?

Ans: The Open System Interconnection (OSI) reference model describes how information from a software application in one computer moves through a network medium to a software application in another computer. The OSI reference model is a conceptual model composed of seven layers, each specifying particular network functions. The model was developed by the International Standardization Organization (ISO) in 1984, and it is now considered the primary architectural model for inter-computer communications.

Q-4. What are the seven layers of ISO's OSI model?

Ans:- The seven layers are:

Application Layer

Presentation Layer

Session Layer

Transport Layer

Network Layer

Data Link Layer

Physical Layer

Q-5. Briefly write functionalities of different OSI layers?

Ans: The OSI Reference Model includes seven layers. Basic functionality of each of them is as follows:

7. *Application Layer:* Provides Applications with access to network services.

6. *Presentation Layer:* Determines the format used to exchange data among networked computers.

5. *Session Layer:* Allows two applications to establish, use and disconnect a connection between them called a session. Provides for name recognition and additional functions like security, which are needed to allow applications to communicate over the network.

4. *Transport Layer:* Ensures that data is delivered error free, in sequence and with no loss, duplications or corruption. This layer also repackages data by assembling long messages into lots of smaller messages for sending, and repackaging the smaller messages into the original larger message at the receiving end.

3. *Network Layer:* This is responsible for addressing messages and data so they are sent to the correct destination, and for translating logical addresses and names (like a machine name FLAME) into physical addresses. This layer is also responsible for finding a path through the network to the destination computer.

2. *Data-Link Layer:* This layer takes the data frames or messages from the Network Layer and provides for their actual transmission. At the receiving computer, this layer receives the incoming data and sends it to the network layer for handling. The Data-Link Layer also provides error-free delivery of data between the two computers by using the physical layer. It does this by packaging the data from the Network Layer into a frame, which includes error detection information. At the receiving computer, the Data-Link Layer reads the incoming frame, and generates its own error detection information based on the received frames data. After receiving the entire frame, it then compares its error detection value with that of the incoming frames, and if they match, the frame has been received correctly.

1. *Physical Layer*: Controls the transmission of the actual data onto the network cable. It defines the electrical signals, line states and encoding of the data and the connector types used. An example is 10BaseT.

Q-6. How two adjacent layers communicate in a layered network? (or What do you mean by Service Access Point?)

Ans: In layered network, each layer has various entities and entities of layer i provide service to the entities of layer i+1. The services can be accessed through service access

point (SAP), which has some address through which the layer i+1 will access the services provided by layer i.

Q-7. What are the key functions of data link layer?

Ans: Data link layer transfers data in a structured and reliable manner so that the service provided by the physical layer is utilized by data link layer. Main function of data link layer is framing and media access control.

Q8. What do you mean by Protocol?

Ans: In the context of data networking, a **protocol** *is a formal set of rules and conventions that governs how computers exchange information over a network medium*. A protocol implements the functions of one or more of the OSI layers.

2.1.1 Introduction

A simplified model of a data communication system is shown in Fig. 2.1.1. Here there are five basic components:

• **Source**: Source is where the data is originated. Typically it is a computer, but it can be any other electronic equipment such as telephone handset, video camera, etc, which can generate data for transmission to some destination. The data to be sent is represented by x(t).

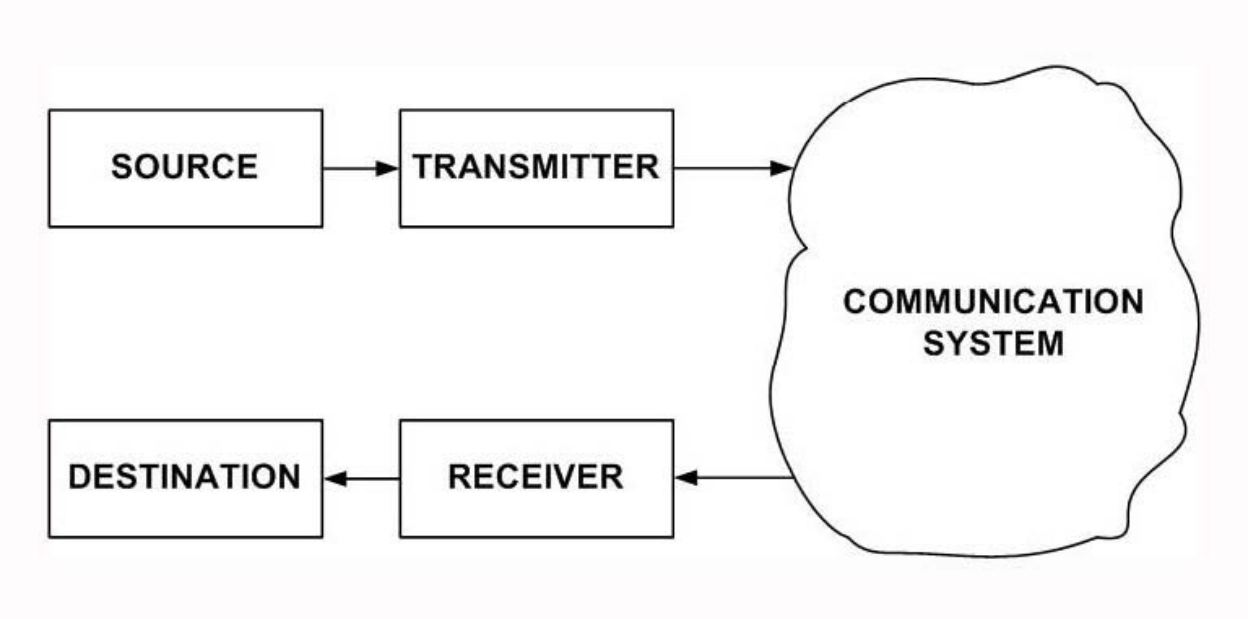


Figure 2.1.1 *Simplified model of a data communication system*

• **Transmitter**: As data cannot be sent in its native form, it is necessary to convert it into signal. This is performed with the help of a transmitter such as modem. The signal that is sent by the transmitter is represented by s(t).

• **Communication Medium**: The signal can be sent to the receiver through a communication medium, which could be a simple twisted-pair of wire, a coaxial cable, optical fiber or wireless communication system. It may be noted that the signal that comes out of the communication medium is s'(t), which is different from s(t) that was sent by the transmitter. This is due to various impairments that the signal suffers as it passes through the communication medium.

• **Receiver**: The receiver receives the signal s'(t) and converts it back to data d'(t) before forwarding to the destination. The data that the destination receives may not be identical to that of d(t), because of the corruption of data.

• **Destination**: Destination is where the data is absorbed. Again, it can be a computer system, a telephone handset, a television set and so on.

2.1.2 Data

Data refers to information that conveys some meaning based on some mutually agreed up rules or conventions between a sender and a receiver and today it comes in a variety of forms such as text, graphics, audio, video and animation.

Data can be of two types; analog and digital. *Analog data* take on continuous values on some interval. Typical examples of analog data are voice and video. The data that are collected from the real world with the help of transducers are continuous-valued or analog in nature. On the contrary, *digital data* take on discrete values. Text or character strings can be considered as examples of digital data. Characters are represented by suitable codes, e.g. ASCII code, where each character is represented by a 7-bit code.

2.1.3 Signal

It is electrical, electronic or optical representation of data, which can be sent over a communication medium. Stated in mathematical terms, a signal is merely a function of the data. For example, a microphone converts voice data into voice signal, which can be sent over a pair of wire. Analog signals are continuous-valued; digital signals are discrete-valued. The independent variable of the signal could be time (speech, for example), space (images), or the integers (denoting the sequencing of letters and numbers in the football score). Figure 2.1.2 shows an analog signal.

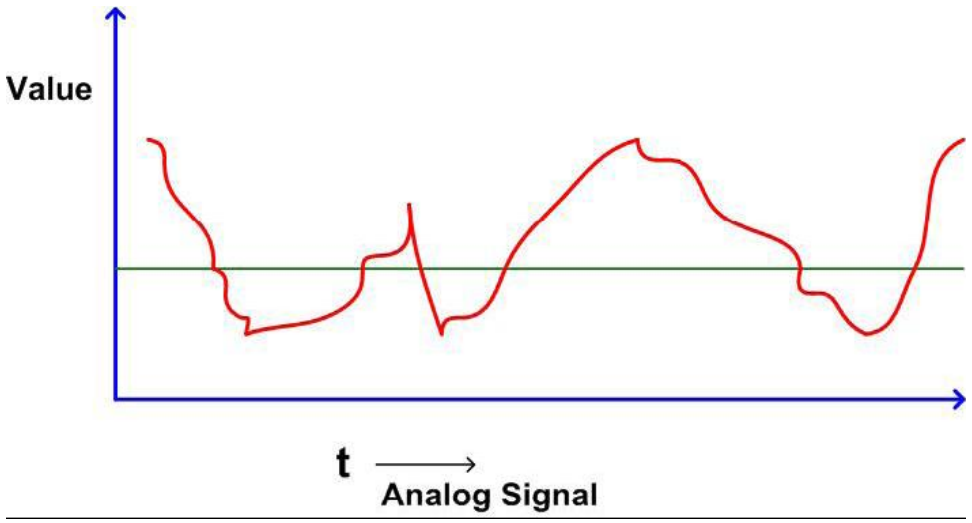


Figure 2.1.2 *Analog signal*

Digital signal can have only a limited number of defined values, usually two values 0 and 1, as shown in Fig. 2.1.3.

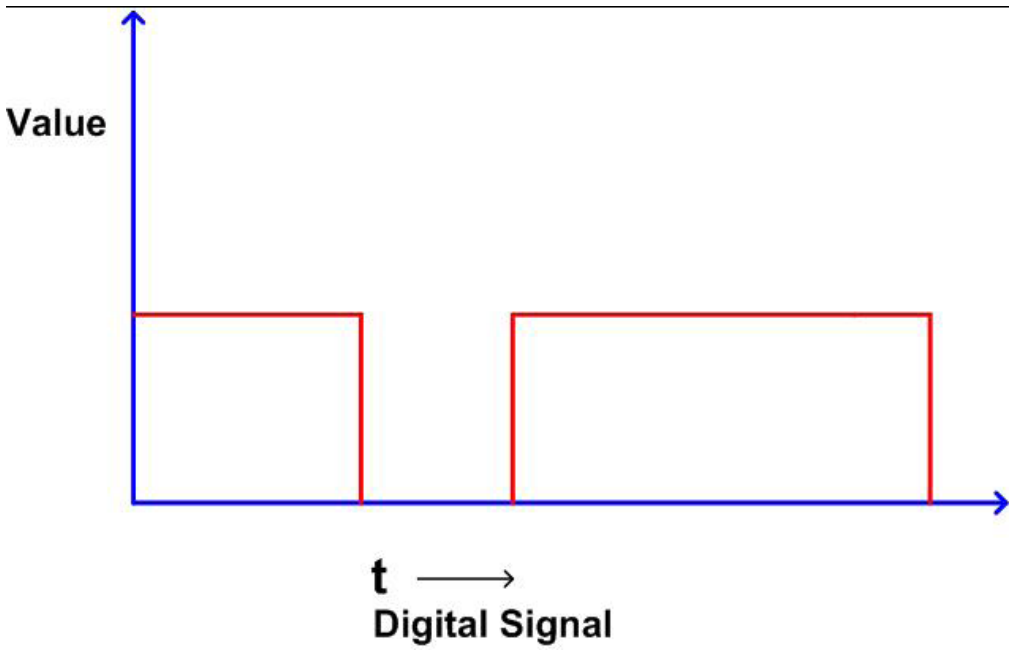
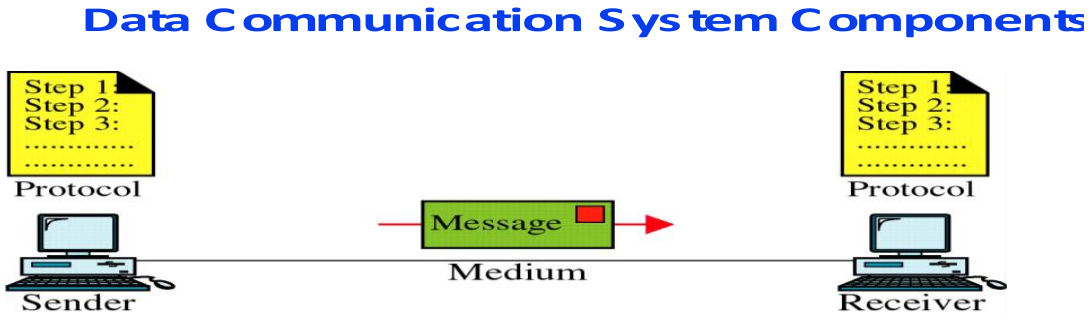


Figure 2.1.3 Digital signal

Signaling: It is an act of sending signal over communication medium

Transmission: Communication of data by propagation and processing is known as transmission. .



Network Technologies

There is no generally accepted taxonomy into which all computer networks fit, but two dimensions stand out as important: **Transmission Technology** and **Scale**. The classifications based on these two basic approaches are considered in this section.

Classification Based on Transmission Technology

Computer networks can be broadly categorized into two types based on transmission technologies:

Broadcast networks

Point-to-point networks

Figure 2-2

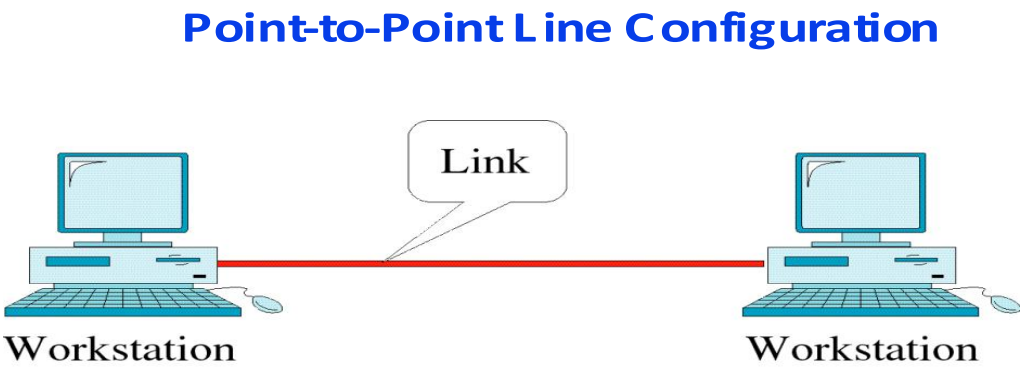
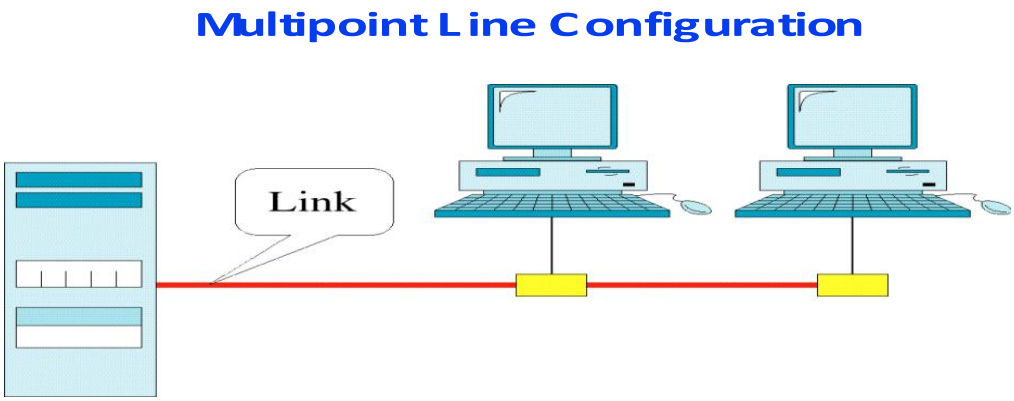


Figure 2-3



Broadcast Networks

Broadcast network have a single communication channel that is shared by all the machines on the network as shown in Figs.1.1.2 and 1.1.3. All the machines on the network receive short messages, called packets in certain contexts, sent by any machine.

An address field within the packet specifies the intended recipient. Upon receiving a packet, machine checks the address field. If packet is intended for itself, it processes the packet; if packet is not intended for itself it is simply ignored. This system generally also allows possibility of addressing the packet to all destinations(all nodes on the network). When such a packet is transmitted and received by all the machines on the network. This mode of operation is known as Broadcast Mode. Some Broadcast systems also supports transmission to a sub-set of machines, something known as Multicasting.

Point-to-Point Networks

A network based on point-to-point communication is shown in Fig. 1.1.4. The end devices that wish to communicate are called stations. The switching devices are called nodes. Some Nodes connect to other nodes and some to attached stations. It uses FDM or TDM for node-to-node communication. There may exist multiple paths between a source-destination pair for better network reliability. The switching nodes are not concerned with the contents of data. Their purpose is to provide a switching facility that will move data from node to node until they reach the destination. As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks normally use are point-to-point communication.

Figure 2-5

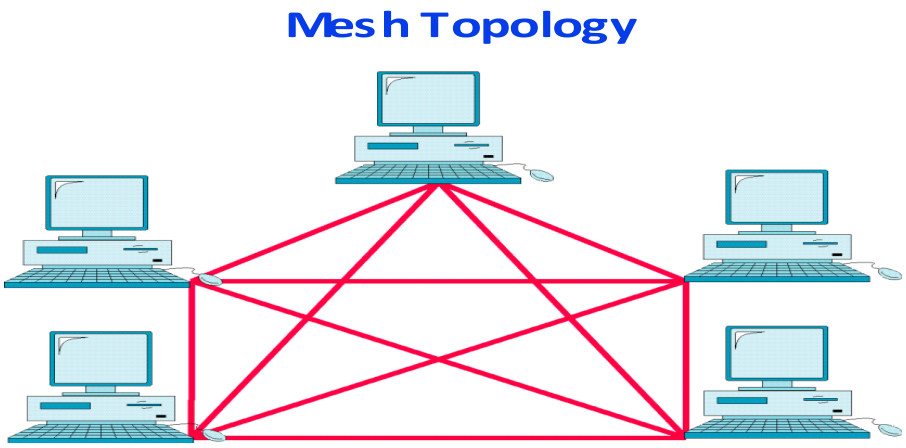


Figure 2-6

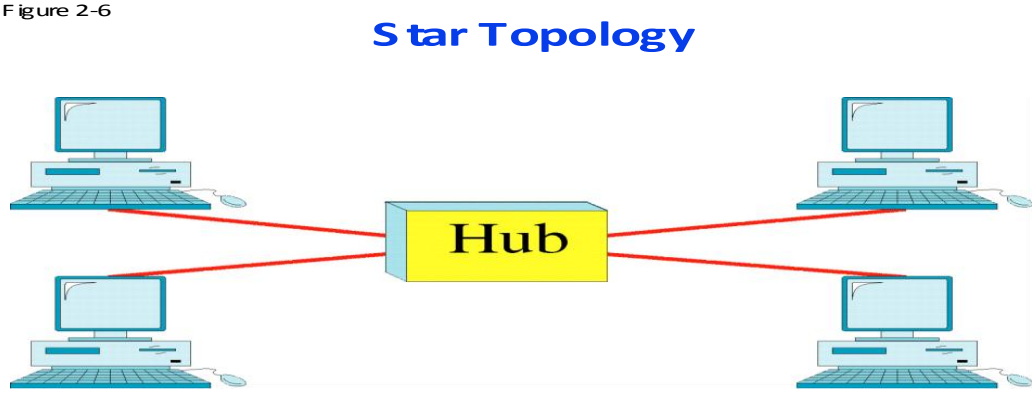
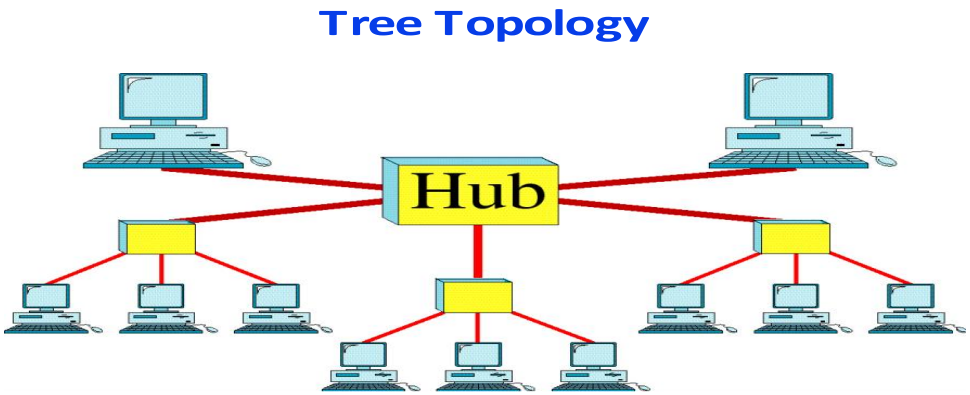


Figure 2-7



PROTOCOL

The OSI model provides a conceptual framework for communication between computers, but the model itself is not a method of communication. Actual communication is made possible by using communication protocols. In the context of data networking, a protocol is a formal set of rules and conventions that governs how computers exchange information over a network medium. A protocol implements the functions of one or more of the OSI layers.

A wide variety of communication protocols exist. Some of these protocols include LAN protocols, WAN protocols, network protocols, and routing protocols. LAN protocols operate at the physical and data link layers of the OSI model and define communication over various LAN media. WAN protocols operate at the lowest three layers of the OSI model and define communication over the various wide-area media. Routing protocols are network layer protocols that are responsible for exchanging information between

routers so that the routers can select the proper path for network traffic. Finally, network protocols are the various upper-layer protocols that exist in a given protocol suite. Many protocols rely on others for operation. For example, many routing protocols use network protocols to exchange information between routers. This concept of building upon the layers already in existence is the foundation of the OSI model.

OSI Model and Communication between Systems

Information being transferred from a software application in one computer system to a software application in another must pass through the OSI layers. For example, if a software application in System A has information to transmit to a software application in System B, the application program in System A will pass its information to the application layer (Layer 7) of System A. The application layer then passes the information to the presentation layer (Layer 6), which relays the data to the session layer (Layer 5), and so on down to the physical layer (Layer 1). At the physical layer, the information is placed on the physical network medium and is sent across the medium to System B. The physical layer of System B removes the information from the physical medium, and then its physical layer passes the information up to the data link layer (Layer 2), which passes it to the network layer (Layer 3), and so on, until it reaches the application layer (Layer 7) of System B. Finally, the application layer of System B passes the information to the recipient application program to complete the communication process.

Interaction between OSI Model Layers

A given layer in the OSI model generally communicates with three other OSI layers: the layer directly above it, the layer directly below it, and its peer layer in other networked computer systems. The data link layer in System A, for example, communicates with the network layer of System A, the physical layer of System A, and the data link layer in System B. Figure1.2.3 illustrates this example.

Services and service access points

One OSI layer communicates with another layer to make use of the services provided by the second layer. The services provided by adjacent layers help a given OSI layer communicate with its peer layer in other computer systems. Three basic elements are involved in layer services: the service user, the service provider, and the service access point (SAP).

In this context, the service user is the OSI layer that requests services from an adjacent OSI layer. The service provider is the OSI layer that provides services to service users. OSI layers can provide services to multiple service users. The SAP is a conceptual location at which one OSI layer can request the services of another OSI layer.

OSI Model Layers and Information Exchange

The seven OSI layers use various forms of control information to communicate with their peer layers in other computer systems. This control information consists of specific requests and instructions that are exchanged between peer OSI layers.

Control information typically takes one of two forms: headers and trailers. Headers are pretended to data that has been passed down from upper layers. Trailers are appended to data that has been passed down from upper layers. An OSI layer is not required to attach a header or a trailer to data from upper layers.

Headers, trailers, and data are relative concepts, depending on the layer that analyzes the information unit. At the network layer, for example, an information unit consists of a Layer 3 header and data. At the data link layer, however, all the information passed down by the network layer (the Layer 3 header and the data) is treated as data.

In other words, the data portion of an information unit at a given OSI layer potentially can contain headers, trailers, and data from all the higher layers. This is known as encapsulation. Figure 1-6 shows how the header and data from one layer are encapsulated into the header of the next lowest layer.

Information Exchange Process

The information exchange process occurs between peer OSI layers. Each layer in the source system adds control information to data, and each layer in the destination system analyzes and removes the control information from that data.

If system A has data from software application to send to System B, the data is passed to the application layer. The application layer in System A then communicates any control information required by the application layer in System B by pre-pending a header to the data. The resulting information unit (a header and the data) is passed to the presentation layer, which pre-pends its own header containing control information intended for the presentation layer in System B. The information unit grows in size as each layer pre-pends its own header (and, in some cases, a trailer) that contains control information to be used by its peer layer in System B. At the physical layer, the entire information unit is placed onto the network medium.

The physical layer in System B receives the information unit and passes it to the data link layer. The data link layer in System B then reads the control information contained in the header pre-pended by the data link layer in System A. The header is then removed, and the remainder of the information unit is passed to the network layer. Each layer performs the same actions: The layer reads the header from its peer layer, strips it off, and passes the remaining information unit to the next highest layer. After the application layer performs these actions, the data is passed to the recipient software application in System B, in exactly the form in which it was transmitted by the application in System A

TransmissionMedia

Introduction

Transmission media can be defined as physical path between transmitter and receiver in a data transmission system. And it may be classified into two types as shown in Fig. 2.2.1.

Guided: Transmission capacity depends critically on the medium, the length, and whether the medium is point-to-point or multipoint (e.g. LAN). Examples are co-axial cable, twisted pair, and optical fiber.
Unguided: provides a means for transmitting electro-magnetic signals but do not guide them. Example wireless transmission.

Characteristics and quality of data transmission are determined by medium and signal characteristics. For guided media, the medium is more important in determining the limitations of transmission. While in case of unguided media, the bandwidth of the signal produced by the transmitting antenna and the size of the antenna is more important than the medium. Signals at lower frequencies are omni-directional (propagate in all directions). For higher frequencies, focusing the signals into a directional beam is possible. These properties determine what kind of media one should use in a particular application. In this lesson we shall discuss the characteristics of various transmission media, both guided and unguided.

Guided transmission media

In this section we shall discuss about the most commonly used guided transmission media such as twisted-pair of cable, coaxial cable and optical fiber.

Twisted Pair

In twisted pair technology, two copper wires are strung between two points:

• The two wires are typically ``twisted" together in a helix to reduce interference between the two conductors .Twisting decreases the cross-talk interference between adjacent pairs in a cable. Typically, a number of pairs are bundled together into a cable by wrapping them in a tough protective sheath. Actually, they carry only analog signals. However, the ``analog" signals can very closely correspond to the square waves representing bits, so we often think of them as carrying digital data. Data rates of several Mbps common. Spans distances of several kilometers. Data rate determined by wire thickness and length. In addition, shielding to eliminate interference from other wires impacts signal-to-noise ratio, and ultimately, the data rate.

Good, low-cost communication. Indeed, many sites already have twisted pair installed in offices -- existing phone lines!

Typical characteristics: Twisted-pair can be used for both analog and digital communication. The data rate that can be supported over a twisted-pair is inversely proportional to the square of the line length. Maximum transmission distance of 1 Km can be achieved for data rates up to 1 Mb/s. For analog voice signals, amplifiers are required about every 6 Km and for digital signals, repeaters are needed for about 2 Km. To reduce interference, the twisted pair can be shielded with metallic braid. This type of wire is known as Shielded Twisted-Pair (STP) and the other form is known as Unshielded Twisted-Pair (UTP).

Use: The oldest and the most popular use of twisted pair are in telephony. In LAN it is commonly used for point-to-point short distance communication (say, 100m) within a building or a room.

Base band Coaxial With ``coax", the medium consists of a copper core surrounded by insulating material and a braided outer conductor as shown in Fig. 2.2.3. The term base band indicates digital transmission (as opposed to broadband analog).

Physical connection consists of metal pin touching the copper core. There are two common ways to connect to a coaxial cable:

1. With vampire taps, a metal pin is inserted into the copper core. A special tool drills a hole into the cable, removing a small section of the insulation, and a special connector is screwed into the hole. The tap makes contact with the copper core.
2. With a T-junction, the cable is cut in half, and both halves connect to the T-junction. A T-connector is analogous to the signal splitters used to hook up multiple TVs to the same cable wire.

Characteristics: Co-axial cable has superior frequency characteristics compared to twisted-pair and can be used for both analog and digital signaling. In baseband LAN, the data rates lies in the range of 1 KHz to 20 MHz over a distance in the range of 1 Km. Co-axial cables typically have a diameter of 3/8". Coaxial cables are used both for baseband and broadband communication. For broadband CATV application coaxial cable of 1/2" diameter and 75 Ω impedance is used. This cable offers bandwidths of 300 to 400 MHz facilitating high-speed data communication with low bit-error rate. In broadband signaling, signal propagates only in one direction, in contrast to propagation in both directions in baseband signaling. Broadband cabling uses either dual-cable scheme or single-cable scheme with a headend to facilitate flow of signal in one direction. Because of the shielded, concentric construction, co-axial cable is less susceptible to interference and cross talk than the twisted-pair. For long distance communication, repeaters are needed for every kilometer or so. Data rate depends on physical properties of cable, but 10 Mbps is typical.

Use: One of the most popular use of co-axial cable is in cable TV (CATV) for the distribution of TV signals. Another importance use of co-axial cable is in LAN.

Broadband Coaxial

The term broadband refers to analog transmission over coaxial cable. (Note, however, that the telephone folks use broadband to refer to any channel wider than 4 kHz). The technology:

• Typically bandwidth of 300 MHz, total data rate of about 150 Mbps.
•Operates at distances up to 100 km (metropolitan area!).
• Uses analog signaling.
• Technology used in cable television. Thus, it is already available at sites such as universities that may have TV classes.

• Total available spectrum typically divided into smaller channels of 6 MHz each. That is, to get more than 6MHz of bandwidth, you have to use two smaller channels and somehow combine the signals.
• Requires amplifiers to boost signal strength; because amplifiers are one way, data flows in only one direction.

Two types of systems have emerged:

1. Dual cable systems use two cables, one for transmission in each direction:
One cable is used for receiving data. Second cable used to communicate with headend. When a node wishes to transmit data, it sends the data to a special node called the headend. The headend then resends the data on the first cable. Thus, the headend acts as a root of the tree, and all data must be sent to the root for redistribution to the other nodes.
2. Midsplit systems divide the raw channel into two smaller channels, with each sub channel having the same purpose as above. Which is better, broadband or base band? There is rarely a simple answer to such questions. Base band is simple to install, interfaces are inexpensive, but doesn't have the same range. Broadband is more complicated, more expensive, and requires regular adjustment by a trained technician, but offers more services (e.g., it carries audio and video too).

Fiber Optics

In fiber optic technology, the medium consists of a hair-width strand of silicon or glass, and the signal consists of pulses of light. For instance, a pulse of light means ``1'', lack of pulse means ``0''. It has a cylindrical shape and consists of three concentric sections: the core, the cladding, and the jacket as shown in Fig. 2.2.4.

The core, innermost section consists of a single solid dielectric cylinder of diameter d1 and of refractive index n1. The core is surrounded by a solid dielectric cladding of refractive index n2 that is less than n1. As a consequence, the light is propagated through multiple total internal reflection. The core material is usually made of ultra pure fused silica or glass and the cladding is either made of glass or plastic. The cladding is surrounded by a jacket made of plastic. The jacket is used to protect against moisture, abrasion, crushing and other environmental hazards.

Three components are required:

1. Fiber medium: Current technology carries light pulses for tremendous distances (e.g., 100s of kilometers) with virtually no signal loss.
2.Light source: typically a Light Emitting Diode (LED) or laser diode. Running current through the material generates a pulse of light.
3. A photo diode light detector, which converts light pulses into electrical signals.

Advantages:

1. Very high data rate, low error rate. 1000 Mbps (1 Gbps) over distances of kilometers common. Error rates are so low they are almost negligible.
2. Difficult to tap, which makes it hard for unauthorized taps as well. This is responsible for higher reliability of this medium. How difficult is it to prevent coax taps? Very difficult indeed, unless one can keep the entire cable in a locked room!
3. Much thinner (per logical phone line) than existing copper circuits. Because of its thinness, phone companies can replace thick copper wiring with fibers having much more capacity for same volume. This is important because it means that aggregate phone capacity can be upgraded without the need for finding more physical space to hire the new cables.
4. Not susceptible to electrical interference (lightning) or corrosion (rust).
5. Greater repeater distance than coax.

Disadvantages:

• Difficult to tap. It really is point-to-point technology. In contrast, tapping into coax is trivial. No special training or expensive tools or parts are required.
• One-way channel. Two fibers needed to get full duplex (both ways) communication.
Optical Fiber works in three different types of modes (or we can say that we have 3 types of communication using Optical fiber). Optical fibers are available in two varieties; Multi-Mode Fiber (MMF) and Single-Mode Fiber (SMF). For multi-mode fiber the core and cladding diameter lies in the range 50-200μm and 125-400μm, respectively. Whereas in single-mode fiber, the core and cladding diameters lie in the range 8-12μm and 125μm, respectively. Single-mode fibers are also known as Mono-Mode Fiber. Moreover, both single-mode and multi-mode fibers can have two types; step index and graded index. In the former case the refractive index of the core is uniform throughout and at the core cladding boundary there is an abrupt change in refractive index. In the later case, the refractive index of the core varies radially from the centre to the core-cladding boundary from n1 to n2 in a linear manner. Fig. 2.2.5 shows the optical fiber transmission modes.

Figure 2.2.5 Schematics of three optical fiber types, (a) Single-mode step-index, (b) Multi-mode step-index, and (c) Multi-mode graded-index

Characteristics: Optical fiber acts as a dielectric waveguide that operates at optical frequencies (1014 to 1015 Hz). Three frequency bands centered around 850,1300 and 1500 nanometers are used for best results. When light is applied at one end of the optical fiber core, it reaches the other end by means of total internal reflection because of the choice of refractive index of core and cladding material (n1 > n2). The light source can be either light emitting diode (LED) or injection laser diode (ILD). These semiconductor devices emit a beam of light when a voltage is applied across the device. At the receiving end, a photodiode can be used to detect the signal-encoded light. Either PIN detector or APD (Avalanche photodiode) detector can be used as the light detector.

In a multi-mode fiber, the quality of signal-encoded light deteriorates more rapidly than single-mode fiber, because of interference of many light rays. As a consequence, single-

mode fiber allows longer distances without repeater. For multi-mode fiber, the typical maximum length of the cable without a repeater is 2km, whereas for single-mode fiber it is 20km.

Fiber Uses: Because of greater bandwidth (2Gbps), smaller diameter, lighter weight, low attenuation, immunity to electromagnetic interference and longer repeater spacing, optical fiber cables are finding widespread use in long-distance telecommunications. Especially, the single mode fiber is suitable for this purpose. Fiber optic cables are also used in high-speed LAN applications. Multi-mode fiber is commonly used in LAN.

- Long-haul trunks-increasingly common in telephone network (Sprint ads)
- Metropolitan trunks-without repeaters (average 8 miles in length)
- Rural exchange trunks-link towns and villages
- Local loops-direct from central exchange to a subscriber (business or home)
- Local area networks-100Mbps ring networks.

Unguided Transmission

Unguided transmission is used when running a physical cable (either fiber or copper) between two end points is not possible. For example, running wires between buildings is probably not legal if the building is separated by a public street.

Infrared signals typically used for short distances (across the street or within same room),

Microwave signals commonly used for longer distances (10's of km). Sender and receiver use some sort of dish antenna as shown in Fig. 2.2.6.

Difficulties:

1. Weather interferes with signals. For instance, clouds, rain, lightning, etc. may adversely affect communication.
2. Radio transmissions easy to tap. A big concern for companies worried about competitors stealing plans.
3. Signals bouncing off of structures may lead to out-of-phase signals that the receiver must filter out.

Satellite Communication

Satellite communication is based on ideas similar to those used for line-of-sight. A communication satellite is essentially a big microwave repeater or relay station in the sky. Microwave signals from a ground station is picked up by a transponder, amplifies the signal and rebroadcasts it in another frequency, which can be received by ground stations at long distances as shown in Fig. 2.2.7.

To keep the satellite stationary with respect to the ground based stations, the satellite is placed in a geostationary orbit above the equator at an altitude of about 36,000 km. As the spacing between two satellites on the equatorial plane should not be closer than 40, there can be 360/4 = 90 communication satellites in the sky at a time. A satellite can be used for point-to-point communication between two ground-based stations or it can be used to broadcast a signal received from one station to many ground-based

stations as shown in Fig. 2.2.8. Number of geo-synchronous satellites limited (about 90 total, to minimize interference). International agreements regulate how satellites are used, and how frequencies are allocated. Weather affects certain frequencies. Satellite transmission differs from terrestrial communication in another important way: One-way propagation delay is roughly 270 ms. In interactive terms, propagation delay alone inserts a 1 second delay between typing a character and receiving its echo. Characteristics: Optimum frequency range for satellite communication is 1 to 10 GHz. The most popular frequency band is referred to as 4/6 band, which uses 3.7 to 4.2 GHz for down link and 5.925 to 6.425 for uplink transmissions. The 500 MHz bandwidth is usually split over a dozen transponders, each with 36 MHz bandwidth. Each 36 MHz bandwidth is shared by time division multiplexing. As this preferred band is already saturated, the next highest band available is referred to as 12/14 GHz. It uses 14 to 14.5GHz for upward transmission and 11.7 to 12.2 GHz for downward transmissions. Communication satellites have several unique properties. The most important is the long communication delay for the round trip (about 270 ms) because of the long distance (about 72,000 km) the signal has to travel between two earth stations. This poses a number of problems, which are to be tackled for successful and reliable communication.

Another interesting property of satellite communication is its broadcast capability. All stations under the downward beam can receive the transmission. It may be necessary to send encrypted data to protect against piracy.

Use: Now-a-days communication satellites are not only used to handle telephone, telex and television traffic over long distances, but are used to support various internet based services such as e-mail, FTP, World Wide Web (WWW), etc. New types of services, based on communication satellites, are emerging.

Comparison/contrast with other technologies:

1. Propagation delay very high. On LANs, for example, propagation time is in nanoseconds -- essentially negligible.
2. One of few alternatives to phone companies for long distances.
3. Uses broadcast technology over a wide area - everyone on earth could receive a message at the same time!
4. Easy to place unauthorized taps into signal.

Satellites have recently fallen out of favor relative to fiber.

However, fiber has one big disadvantage: no one has it coming into their house or building, whereas anyone can place an antenna on a roof and lease a satellite channel.

Switched Communication Networks

Lesson

1 Switching Techniques: Circuit Switching

Specific Instructional Objectives

At the end of this lesson the student will be able to:

• Understand the need for circuit switching
• Specify the components of a switched communication network
• Explain how circuit switching takes place
• Explain how switching takes place using space-division and time-division switching
• Explain how routing is performed
• Explain how signalling in performed

Introduction

When there are many devices, it is necessary to develop suitable mechanism for communication between any two devices. One alternative is to establish point-to-point communication between each pair of devices using mesh topology. However, mesh topology is impractical for large number of devices, because the number of links increases exponentially (n(n-1)/2, where n is the number of devices) with the number of devices. A better alternative is to use switching techniques leading to switched communication network. In the switched network methodology, the network consists of a set of interconnected nodes, among which information is transmitted from source to destination via different routes, which is controlled by the switching mechanism. A basic model of a switched communication is shown in Fig. 4.1.1. The end devices that wish to communicate with each other are called stations. The switching devices are called nodes. Some nodes connect to other nodes and some are to connected to some stations. Key features of a switched communication network are given below:

• Network Topology is not regular.
• Uses FDM or TDM for node-to-node communication.
• There exist multiple paths between a source-destination pair for better network reliability.
• The switching nodes are not concerned with the contents of data.
    • Their purpose is to provide a switching facility that will move data from node to node until they reach the destination.

The switching performed by different nodes can be categorized into the following three types:

• Circuit Switching
• Packet Switching
• Message Switching

Figure 14-1

Switched Network

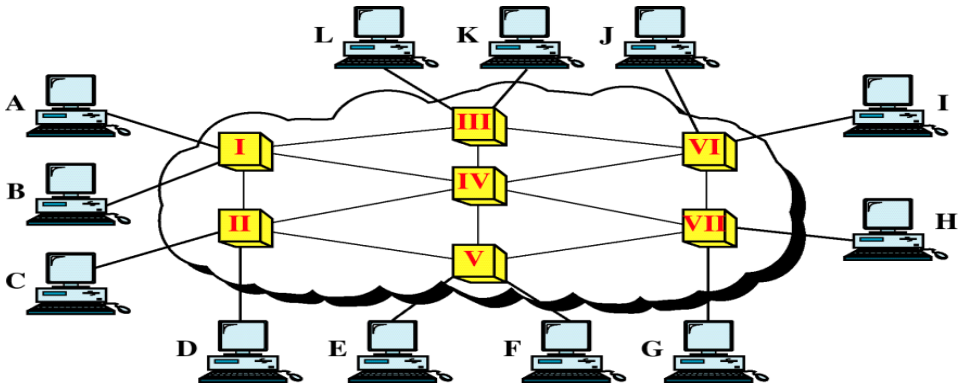


Figure 14-3

Circuit-Switched Network

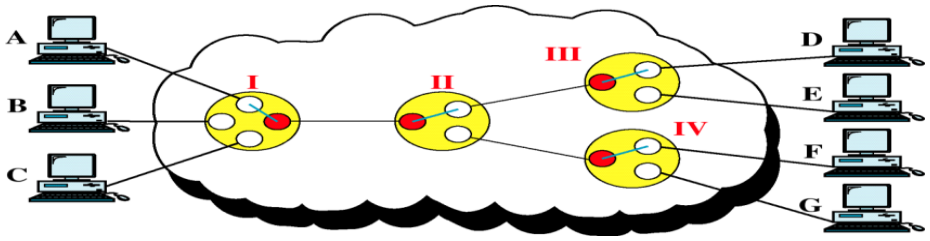


Figure 14-14

Datagram Approach

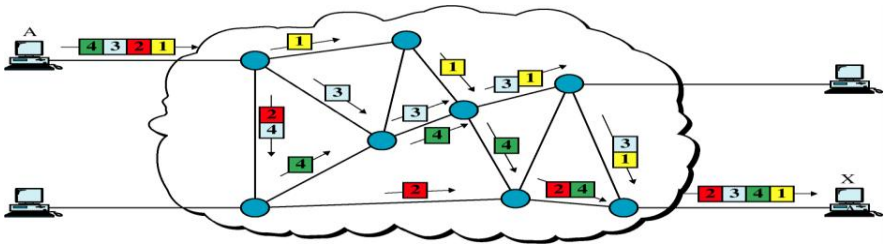
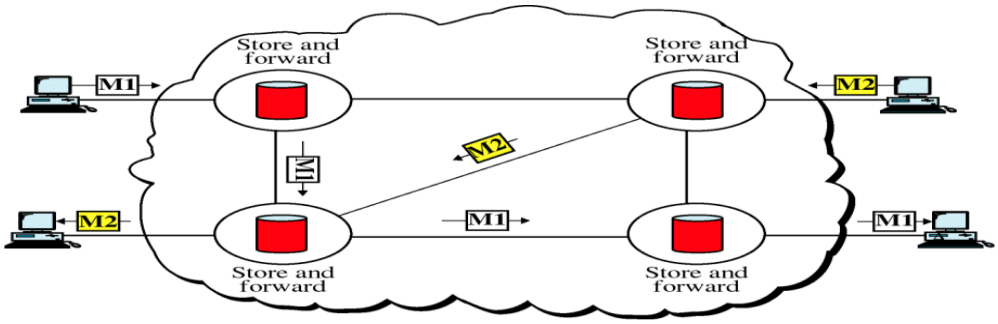


Figure 14-17

Message Switching



Circuit switching Technique

Communication via circuit switching implies that there is a dedicated communication path between the two stations. The path is a connected through a sequence of links between network nodes. On each physical link, a logical channel is dedicated to the connection. Circuit switching is commonly used technique in telephony, where the caller sends a special message with the address of the callee (i.e. by dialling a number) to state its destination. It involved the following three distinct steps,

Circuit Establishment: To establish an end-to-end connection before any transfer of data.

Some segments of the circuit may be a dedicated link, while some other segments may be shared.

Data transfer:

• Transfer data is from the source to the destination.
• The data may be analog or digital, depending on the nature of the network.
• The connection is generally full-duplex.

Circuit disconnect:

• Terminate connection at the end of data transfer.

• Signals must be propagated to deallocate the dedicated resources.

Thus the actual physical electrical path or circuit between the source and destination host must be established before the message is transmitted. This connection, once established, remains exclusive and continuous for the complete duration of information exchange and the circuit becomes disconnected only when the source wants to do so.
Switching Node

Let us consider the operation of a single circuit switched node comprising a collection of stations attached to a central switching unit, which establishes a dedicated path between any two devices that wish to communicate.

Major elements of a single-node network are summarized below:

• Digital switch: That provides a transparent (full-duplex) signal path between any pair of attached devices.
• Network interface: That represents the functions and hardware needed to connect digital devices to the network (like telephones).
• Control unit: That establishes, maintains, and tears down a connection.

The simplified schematic diagram of a switching node is shown in Fig. 4.1.3. An important characteristic of a circuit-switch node is whether it is blocking or non-blocking. A blocking network is one, which may be unable to connect two stations because all possible paths between them are already in use. A non-blocking network permits all stations to be connected (in pairs) at once and grants all possible connection requests as long as the called party is free. For a network that supports only voice traffic, a blocking configuration may be acceptable, since most phone calls are of short duration. For data applications, where a connection may remain active for hours, non-blocking configuration is desirable. Circuit switching uses any of the three technologies: Space-division switches, Time-division switches or a combination of both. In Space-division switching, the paths in the circuit are separated with each other spatially, i.e. different ongoing connections, at a same instant of time, uses different switching paths, which are separated spatially. This was originally developed for the analog environment, and has been carried over to the digital domain. Some of the space switches are crossbar switches, Multi-stage switches (e.g. Omega Switches). A crossbar switch is shown in Fig. 4.1.4. Basic building block of the switch is a metallic crosspoint or semiconductor gate that can be enabled or disabled by a control unit. • The number of crosspoints grows with the square of the number of attached stations.
• Costly for a large switch.
• The failure of a crosspoint prevents connection between the two devices whose lines intersect at that crosspoint.
• The crosspoints are inefficiently utilized.
• Only a small fraction of crosspoints are engaged even if all of the attached devices are active.

Some of the above problems can be overcome with the help of multistage space division switches. By splitting the crossbar switch into smaller units and interconnecting them, it is possible to build multistage switches with fewer crosspoints.

Three-stage space division switch: In this case the number of crosspoints needed goes down from 64 to 40. There is more than one path through the network to connect two endpoints, thereby increasing reliability. Multistage switches may lead to blocking. The problem may be tackled by increasing the number or size of the intermediate switches, which also increases the cost. The blocking feature is illustrated in Fig. 4.1.6. after setting

up connections for 1-to-3 and 2-to-4, the switch cannot establish connections for 3-to-6 and 4-to-5.

Time Division Switching

Both voice and data can be transmitted using digital signals through the same switches. All modern circuit switches use digital time-division multiplexing (TDM) technique for establishing and maintaining circuits. Synchronous TDM allows multiple low-speed bit streams to share a high-speed line. A set of inputs is sampled in a round robin manner. The samples are organized serially into slots (channels) to form a recurring frame of slots. During successive time slots, different I/O pairings are enabled, allowing a number of connections to be carried over the shared bus. To keep up with the input lines, the data rate on the bus must be high enough so that the slots recur sufficiently frequently. For 100 full-duplex lines at 19.200 Kbps, the data rate on the bus must be greater than 1.92 Mbps. The source-destination pairs corresponding to all active connections are stored in the control memory. Thus the slots need not specify the source and destination addresses. Schematic diagram of time division switching. Time-division switching uses time-division multiplexing to achieve switching, i.e. different ongoing connections can use same switching path but at different interleaved time intervals. There are two popular methods of time-division switching namely, Time-Slot Interchange (TSI) and the TDM bus. TSI changes the ordering of the slots based on desired connection and it has a random-access memory to store data and flip the time slots as shown in Fig. 4.1.8. The operation of a TSI is depicted in Fig. 4.1.9. As shown in the figure, writing can be performed in the memory sequentially, but data is read selectively. In TDM bus there are several input and outputs connected to a high-speed bus. During a time slot only one particular output switch is closed, so only one connection at a particular instant of time

Public Switched Telephone Networks

Public switched telephone network (PSTN) is an example of circuit-switched network. It's also known as Plain Old Telephone Service (POTS). The switching centres used for the switching are organised in different levels, namely: Regional offices (class 1), Section offices (class 2), primary offices (class 3), Toll offices (class 4) and finally End offices (class 5). Level 1 is at the highest level and Level 5 is the lowest level. Subscribers or the customers are directly connected to these end offices. And each office is connected directly to a number of offices at a level below and mostly a single office at higher level.

Subscriber Telephones are connected, through Local Loops to end offices (or central offices). A small town may have only one end office, but large cities have several end offices. Many end offices are connected to one Toll office, which are connected to primary offices. Several primary offices are connected to a section office, which normally serves more than one state. All regional offices are connected using mesh topology. Accessing the switching station at the end offices is accomplished through dialling. In the past, telephone featured rotary or pulse dialling, in which digital signals were sent to the end office for each dialled digit. This type of dialling was prone to errors due to inconsistency in humans during dialling. Presently, dialling is accomplished by Touch-Tone technique. In this method the user sends a small burst of frequency called dual tone,

because it is a combination of two frequencies. This combination of frequencies sent depends on the row and column of the pressed pad.

The connections are multiplexed when have to send to a switching office, which is one level up. For example, Different connections will be multiplexed when they are to be forwarded from an end-office to Toll office.

TWO MARKS QUESTIONS

Q-1. Which Technologies of this age had led to the emergence of computer network?
Ans: The technologies are Computer technology and Communication technology with the support of VLSI Technology.

Q-2. What are the two broad classifications under which Networks can be divided?
Ans: All computer networks fit in one of the two dimensions namely,
a). **Transmission Technology**, this focuses on the basic underlying physical network, for e.g. whether the nodes share a communication media or each pair of node has a separate dedicated link.
b). **Scale**, it focuses on the scale of network how large is your network.

Q-3. Mention different categories of computer networks (on the basis of scale) and distinguish one from the other.
Ans: Local Area Network (LAN): It is privately owned communication systems that cover a small area, say a building or a complex of buildings. Length is about 10 meters to few kilometers and operates at a high speed like 10 MBPS to 1000 MBPS. It has very low error rate (1:1011).
Metropolitan Area Network (MAN): It is public or privately owned communication system that typically covers a complete city. Speed is about 10 MBPS and follows DQDB (Distributed Queue Double Bus) standard. Its reliability is moderate.
Wide Area Network (WAN): It covers a large geographical area and usually owned by a state. Data transfer rate is low (few KBPS to 10 MBPS) and error rate is much higher.

Q-4. What are the two types of Transmission technologies, basis on which computer networks can be categorized?
Ans: Broadly there are two types of transmission technology:

1**. Broadcast networks**: a single communication channel that is shared by all the machines on the network

2. **Point-to-point networks**: This network consists of many connections between individual pairs of machines. To go from the source to destination a message (or packet) may have to visit one or more intermediate machines

Q-5. What is Internet?

Ans: Internet is a collection of networks or network of networks. Various networks such as LAN and WAN connected through suitable hardware and software to work in a seamless manner. It allows various applications such as e-mail; file transfer, remote log-in, World Wide Web, Multimedia, etc run across the internet.

Q-6. How do you account for higher reliability and scalability of computer network?
Ans: Computer network can have a large number of computers, which can share software, database and other resources. In the event of failure of one computer, its workload can be taken over by other computers. So, it provides higher reliability than centralized computing system.
Requirement of software, hardware, database etc. increases gradually. In centralized computing system, if one computer is not able to serve the purpose, we have to replace it by new one. Replacement of new computer requires lot of investment and effort, which can be avoided in computer network system. If there is need for more, one can buy another powerful computer, add it to computer network and use it. The various resources like computers, peripherals, etc. can be added in a scalable manner.

Q-7. Mention important benefits of computer network.
Ans: Important benefits of computer networks are:

i) Resource sharing
ii)Powerful communication medium
iii) Higher reliability
iv) Higher flexibility
v) Lower cost
vi) Incremental expansion

Q-8. What are the main categories based on which applications of computer network can be categorized?
Ans: The main areas under which the applications for computer network can be categorized are as follows:
Scientific and Technical Computing

– Client Server Model, Distributed Processing
– Parallel Processing, Communication Media

Commercial

– Advertisement, Telemarketing, Teleconferencing
– Worldwide Financial Services

Network for the People (this is the most widely used application nowadays)

– Telemedicine, Distance Education, Access to Remote Information, Person-to-Person Communication, Interactive Entertainment

Q-9 How is computer networks used in marketing and sales, financial services, teleconferencing?

Ans: Computer network have led to a new age of all of these services. They have helped in the following way to individual sector:

Marketing and sales: Computer networks are used extensively in both marketing and sales organizations. Marketing professionals use them to collect, exchange, and analyze data related to customer needs and product development cycles. Sales application includes teleshopping, which uses order-entry computers or telephones connected to order processing network, and online-reservation services for hotels, airlines and so on.

Financial services: Today's financial services are totally depended on computer networks. Application includes credit history searches, foreign exchange and investment services, and electronic fund transfer, which allow user to transfer money without going into a bank (an automated teller machine is an example of electronic fund transfer, automatic pay-check is another).

Teleconferencing: Teleconferencing allows conference to occur without the participants being in the same place. Applications include simple text conferencing (where participants communicate through their normal keyboards and monitor) and video conferencing where participants can even see as well as talk to other fellow participants. Different types of equipments are used for video conferencing depending on what quality of the motion you want to capture (whether you want just to see the face of other fellow participants or do you want to see the exact facial expression).

----------x----------x----------

UNIT II

DATA LINK LAYER

Data link control: Framing – Flow and error control –Protocols for Noiseless and Noisy hannels – HDLC Multiple access: Random access – Controlled access Wired LANS : Ethernet – IEEE standards – standard Ethernet – changes in the standard – Fast Ethernet – Gigabit Ethernet. Wireless LANS : IEEE 802.11–Bluetooth. Connecting LANS: Connecting devices - Backbone networks - Virtual LANS Virtual circuit networks: Architecture and Layers of Frame Relay and ATM.

2.1. Flow Control and Error Control

Introduction

As we have mentioned earlier, for reliable and efficient data communication a great deal of coordination is necessary between at least two machines. Some of these are necessary because of the following constraints:

• Both sender and receiver have limited speed
• Both sender and receiver have limited memory

It is necessary to satisfy the following requirements:

• A fast sender should not overwhelm a slow receiver, which must perform a certain amount of processing before passing the data on to the higher-level software.
• If error occur during transmission, it is necessary to devise mechanism to correct it

The most important functions of Data Link layer to satisfy the above requirements are **error control** and **flow control**. Collectively, these functions are known as **data link control**, as discussed in this lesson.

Flow Control is a technique so that transmitter and receiver with different speed characteristics can communicate with each other. Flow control ensures that a transmitting

station, such as a server with higher processing capability, does not overwhelm a receiving station, such as a desktop system, with lesser processing capability. This is where there is an orderly flow of transmitted data between the source and the destination. **Error Control** involves both error detection and error correction. It is necessary because errors are inevitable in data communication, in spite of the use of better equipment and reliable transmission media based on the current technology. In the preceding lesson we have already discussed how errors can be detected. In this lesson we shall discuss how error control is performed based on retransmission of the corrupted data. When an error is detected, the receiver can have the specified frame retransmitted by the sender. This process is commonly known as **Automatic Repeat Request (ARQ).** For example, Internet's Unreliable Delivery Model allows packets to be discarded if network resources are not available, and demands that ARQ protocols make provisions for retransmission.

Flow Control

Modern data networks are designed to support a diverse range of hosts and communication mediums. Consider a 933 MHz Pentium-based host transmitting data to a 90 MHz 80486/SX. Obviously, the Pentium will be able to drown the slower processor with data. Likewise, consider two hosts, each using an Ethernet LAN, but with the two Ethernets connected by a 56 Kbps modem link. If one host begins transmitting to the other at Ethernet speeds, the modem link will quickly become overwhelmed. In both cases, flow control is needed to pace the data transfer at an acceptable speed.

Flow Control is a set of procedures that tells the sender how much data it can transmit before it must wait for an acknowledgment from the receiver. The flow of data should not be allowed to overwhelm the receiver. Receiver should also be able to inform the transmitter before its limits (this limit may be amount of memory used to store the incoming data or the processing power at the receiver end) are reached and the sender must send fewer frames. Hence, Flow control refers to the set of procedures used to restrict the amount of data the transmitter can send before waiting for acknowledgment.

There are two methods developed for flow control namely Stop-and-wait and Sliding-window. Stop-and-wait is also known as Request/reply sometimes. Request/reply (Stop-and-wait) flow control requires each data packet to be acknowledged by the remote host before the next packet is sent. This is discussed in detail in the following subsection. Sliding window algorithms, used by TCP, permit multiple data packets to be in simultaneous transit, making more efficient use of network Stop-and-Wait

This is the simplest form of flow control where a sender transmits a data frame. After receiving the frame, the receiver indicates its willingness to accept another frame by sending back an ACK frame acknowledging the frame just received. The sender must wait until it receives the ACK frame before sending the next data frame.This is sometimes referred to as ping-pong behavior, request/reply is simple to understand and easy to implement, but not very efficient. In LAN environment with fast links, this isn't much of a concern, but WAN links will spend most of their time idle, especially if several hops are required.

The blue arrows show the sequence of data frames being sent across the link from the sender (top to the receiver (bottom). The protocol relies on two-way transmission (full duplex or half duplex) to allow the receiver at the remote node to return frames acknowledging the successful transmission. The acknowledgements are shown in green in the diagram, and flow back to the original sender. A small processing delay may be introduced between reception of the last byte of a Data PDU and generation of the corresponding ACK.

Major drawback of Stop-and-Wait Flow Control is that only one frame can be in transmission at a time, this leads to inefficiency if propagation delay is much longer than the transmission delay.

Stop-and Wait protocol	Some protocols pretty much require stop-and-wait behavior. For example, Internet's Remote Procedure Call (RPC) Protocol is used to implement subroutine calls from a program on one machine to library routines on another machine. Since most programs are single threaded, the sender has little choice but to wait for a reply before continuing the program and possibly sending another request.

Link Utilization in Stop-and-Wait Let us assume the following:

Transmission time: The time it takes for a station to transmit a frame (normalized to a value of 1).

Propagation delay: The time it takes for a bit to travel from sender to receiver (expressed as a).

a < 1 :The frame is sufficiently long such that the first bits of the frame arrive at the destination before the source has completed transmission of the frame.
– a > 1: Sender completes transmission of the entire frame before the leading bits of the frame arrive at the receiver.
– The link utilization U = 1/(1+2a),
a = Propagation time / transmission time

It is evident from the above equation that the link utilization is strongly dependent on the ratio of the propagation time to the transmission time. When the propagation time is small, as in case of LAN environment, the link utilization is good. But, in case of long propagation delays, as in case of satellite communication, the utilization can be very poor. To improve the link utilization, we can use the following (sliding-window) protocol instead of using stop-and-wait protocol.

2.2. Sliding Window

With the use of multiple frames for a single message, the stop-and-wait protocol does not perform well. Only one frame at a time can be in transit. In stop-and-wait flow control, if a > 1, serious inefficiencies result. Efficiency can be greatly improved by allowing multiple frames to be in transit at the same time. Efficiency can also be improved by making use of the full-duplex line. To keep track of the frames, sender station sends sequentially numbered frames. Since the sequence number to be used occupies a field in the frame, it should be of limited size. If the header of the frame allows k bits, the

sequence numbers range from 0 to 2k – 1. Sender maintains a list of sequence numbers that it is allowed to send (sender window). The size of the sender's window is at most 2k – 1. The sender is provided with a buffer equal to the window size. Receiver also maintains a window of size 2k – 1. The receiver acknowledges a frame by sending an ACK frame that includes the sequence number of the next frame expected. This also explicitly announces that it is prepared to receive the next N frames, beginning with the number specified. This scheme can be used to acknowledge multiple frames. It could receive frames 2, 3, 4 but withhold ACK until frame 4 has arrived. By returning an ACK with sequence number 5, it acknowledges frames 2, 3, 4 in one go. The receiver needs a buffer of size 1.

Sliding window algorithm is a method of flow control for network data transfers. TCP, the Internet's stream transfer protocol, uses a sliding window algorithm.

A sliding window algorithm places a buffer between the application program and the network data flow. For TCP, the buffer is typically in the operating system kernel, but this is more of an implementation detail than a hard-and-fast requirement.	Buffer in sliding window

Data received from the network is stored in the buffer, from where the application can read at its own pace. As the application reads data, buffer space is freed up to accept more input from the network. The window is the amount of data that can be "read ahead" - the size of the buffer, less the amount of valid data stored in it. Window announcements are used to inform the remote host of the current window size.

Sender sliding Window:

 At any instant, the sender is permitted to send frames with sequencenumbers in a certain range (the sending window)

Receiver sliding Window:

 The receiver always maintains a window of size 1 as shown in It looks for a specific frame (frame 4 as shown in the figure) to arrive in a specific order. If it receives any other frame (out of order), it is discarded and it needs to be resent. However, the receiver window also slides by one as the specific frame is received and accepted as shown in the figure. The receiver acknowledges a frame by sending an ACK frame that includes the sequence number of the next frame expected. This also explicitly announces that it is prepared to receive the next N frames, beginning with the number specified. This scheme can be used to acknowledge multiple frames. It could receive frames 2, 3, 4 but

withhold ACK until frame 4 has arrived. By returning an ACK with sequence number 5, it acknowledges frames 2, 3, 4 at one time. The receiver needs a buffer of size 1.

Receiver sliding window

On the other hand, if the local application can process data at the rate it's being transferred; sliding window still gives us an advantage. If the window size is larger than the packet size, then multiple packets can be outstanding in the network, since the sender knows that buffer space is available on the receiver to hold all of them. Ideally, a steady-state condition can be reached where a series of packets (in the forward direction) and window announcements (in the reverse direction) are constantly in transit. As each new window announcement is received by the sender, more data packets are transmitted. As the application reads data from the buffer (remember, we're assuming the application can keep up with the network), more window announcements are generated. Keeping a series of data packets in transit ensures the efficient use of network resources.

The link utilization in case of Sliding Window Protocol

$U = 1$, for $N > 2a + 1$

$N/(1+2a)$, for $N < 2a + 1$

Where N = the window size,

and a = Propagation time / transmission time

2.3. Error Control Techniques

When an error is detected in a message, the receiver sends a request to the transmitter to retransmit the ill-fated message or packet. The most popular retransmission scheme is known as Automatic-Repeat-Request (ARQ). Such schemes, where receiver asks transmitter to re-transmit if it detects an error, are known as reverse error correction techniques.

Stop-and-Wait ARQ

In Stop-and-Wait ARQ, which is simplest among all protocols, the sender (say station A) transmits a frame and then waits till it receives positive acknowledgement (ACK) or negative acknowledgement (NACK) from the receiver (say station B). Station B sends an ACK if the frame is received correctly, otherwise it sends NACK. Station A sends a new frame after receiving ACK; otherwise it retransmits the old frame, if it receives a NACK.
Stop-And-Wait ARQ technique

To tackle the problem of a lost or damaged frame, the sender is equipped with a timer. In case of a lost ACK, the sender transmits the old frame. In the Fig. 3.3.7, the second PDU of Data is lost during transmission. The sender is unaware of this loss, but starts a timer after sending each PDU.

In this case no ACK is received, and the timer counts down to zero and triggers retransmission of the same PDU by the sender. The sender always starts a timer following transmission, but in the second transmission receives an ACK PDU before the timer expires, finally indicating that the data has now been received by the remote node.
Retransmission due to lost frame

The receiver now can identify that it has received a duplicate frame from the label of the frame and it is discarded

To tackle the problem of damaged frames, say a frame that has been corrupted during the transmission due to noise, there is a concept of NACK frames, i.e. Negative Acknowledge frames. Receiver transmits a NACK frame to the sender if it founds the received frame to be corrupted. When a NACK is received by a transmitter before the time-out, the old frame is sent again
Retransmission due to damaged frame

The main advantage of stop-and-wait ARQ is its simplicity. It also requires minimum buffer size. However, it makes highly inefficient use of communication links, particularly when 'a' is large.

2.4. Go-back-N ARQ

The most popular ARQ protocol is the go-back-N ARQ, where the sender sends the frames continuously without waiting for acknowledgement. That is why it is also called as continuous ARQ. As the receiver receives the frames, it keeps on sending ACKs or a NACK, in case a frame is incorrectly received. When the sender receives a NACK, it retransmits the frame in error plus all the succeeding frames as shown in Fig.3.3.9. Hence, the name of the protocol is go-back-N ARQ. If a frame is lost, the receiver sends NAK after receiving the next frame as shown in Fig. 3.3.10. In case there is long delay before sending the NAK, the sender will resend the lost frame after its timer times out. If the ACK frame sent by the receiver is lost, the sender resends the frames after its timer times out Assuming full-duplex transmission, the receiving end sends piggybacked acknowledgement by using some number in the ACK field of its data frame. Let us assume that a 3-bit sequence number is used and suppose that a station sends frame 0 and gets back an RR1, and then sends frames 1, 2, 3, 4, 5, 6, 7, 0 and gets another RR1.This might either mean that RR1 is a cumulative ACK or all 8 frames were damaged. This ambiguity can be overcome if the maximum window size is limited to 7, i.e. for a k-bit sequence number field it is limited to 2k-1. The number N (=2k-1) specifies how many frames can be sent without receiving acknowledgement.

If no acknowledgement is received after sending N frames, the sender takes the help of a timer. After the time-out, it resumes retransmission. The go-back-N protocol
also takes care of damaged frames and damaged ACKs. This scheme is little more complex than the previous one but gives much higher throughput.
Assuming full-duplex transmission, the receiving end sends piggybacked acknowledgement by using some number in the ACK field of its data frame. Let us assume that a 3-bit sequence number is used and suppose that a station sends frame 0 and

gets back an RR1, and then sends frames 1, 2, 3, 4, 5, 6, 7, 0 and gets another RR1.This might either mean that RR1 is a cumulative ACK or all 8 frames were damaged. This ambiguity can be overcome if the maximum window size is limited to 7, i.e. for a k-bit sequence number field it is limited to 2k-1. The number N (=2k-1) specifies how many frames can be sent without receiving acknowledgement. If no acknowledgement is received after sending N frames, the sender takes the help of a timer. After the time-out, it resumes retransmission. The go-back-N protocol also takes care of damaged frames and damaged ACKs. This scheme is little more complex than the previous one but gives much higher throughput.

2.5. Selective-Repeat ARQ

The selective-repetitive ARQ scheme retransmits only those for which NAKs are received or for which timer has expired, This is the most efficient among the ARQ schemes, but the sender must be more complex so that it can send out-of-order frames. The receiver also must have storage space to store the post-NAK frames and processing power to reinsert frames in proper sequence.

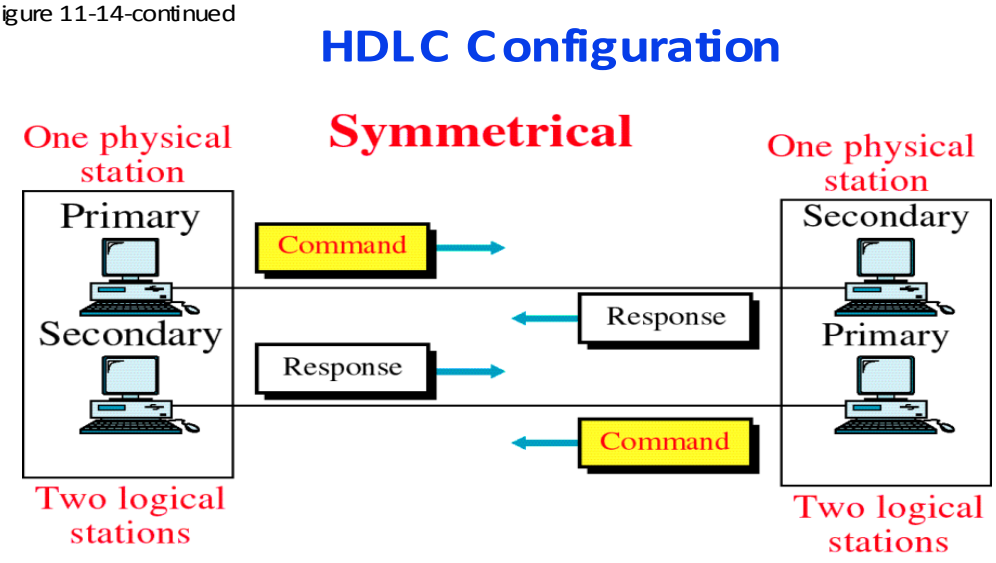
2.6. HDLC

Introduction

HDLC is a bit-oriented protocol. It was developed by the International Organization for Standardization (ISO). It falls under the ISO standards ISO 3309 and ISO 4335. It specifies a packitization standard for serial links. It has found itself being used throughout the world. It has been so widely implemented because it supports both half-duplex and full-duplex communication lines, point-to-point (peer to peer) and multi-point networks, and switched or non-switched channels. HDLC supports several modes of operation, including a simple sliding-window mode for reliable delivery. Since Internet provides retransmission at higher levels (i.e., TCP), most Internet applications use HDLC's unreliable delivery mode, Unnumbered Information.

Other benefits of HDLC are that the control information is always in the same position, and specific bit patterns used for control differ dramatically from those in representing data, which reduces the chance of errors. It has also led to many subsets. Two subsets widely in use are Synchronous Data Link Control (SDLC) and Link Access Procedure-Balanced (LAP-B).

Figure 11-14-continued



In this lesson we shall consider the following aspects of HDLC:

- Stations and Configurations
- Operational Modes
- Non-Operational Modes
- Frame Structure
- Commands and Responses
- HDLC Subsets (SDLC and LAPB)

HDLC Stations and Configurations

HDLC specifies the following three types of stations for data link control:

- Primary Station
- Secondary Station
- Combined Station

Primary Station

Within a network using HDLC as its data link protocol, if a configuration is used in which there is a primary station, it is used as the controlling station on the link. It has the responsibility of controlling all other stations on the link (usually secondary stations). A primary issues commands and secondary issues responses. Despite this important aspect of being on the link, the primary station is also responsible for the organization of data flow on the link. It also takes care of error recovery at the data link level (layer 2 of the OSI model).

Secondary Station

If the data link protocol being used is HDLC, and a primary station is present, a secondary station must also be present on the data link. The secondary station is under the control of the primary station. It has no ability, or direct responsibility for controlling the link. It is only activated when requested by the primary station. It only responds to the primary station. The secondary station's frames are called responses. It can only send response frames when requested by the primary station. A primary station maintains a separate logical link with each secondary station.

Combined Station

A combined station is a combination of a primary and secondary station. On the link, all combined stations are able to send and receive commands and responses without any permission from any other stations on the link. Each combined station is in full control of itself, and does not rely on any other stations on the link. No other stations can control any combined station. HDLC also defines three types of configurations for the three types of stations. The word configuration refers to the relationship between the hardware devices on a link. Following are the three configurations defined by HDLC:

• Unbalanced Configuration
• Balanced Configuration
• Symmetrical Configuration

Unbalanced Configuration

The unbalanced configuration in an HDLC link consists of a primary station and one or more secondary stations. The unbalanced condition arises because one station controls the other stations. In an unbalanced configuration, any of the following can be used:
• Full-Duplex or Half-Duplex operation
• Point to Point or Multi-point networks

Balanced Configuration

The balanced configuration in an HDLC link consists of two or more combined stations. Each of the stations has equal and complimentary responsibility compared to each other. Balanced configurations can use only the following:

• Full - Duplex or Half - Duplex operation
• Point to Point networks
Symmetrical Configuration

This third type of configuration is not widely in use today. It consists of two independent point-to-point, unbalanced station configurations. In this configuration, each station has a primary and secondary status. Each station is logically considered as two stations.

HDLC Operational Modes

A mode in HDLC is the relationship between two devices involved in an exchange; the mode describes who controls the link. Exchanges over unbalanced configurations are always conducted in normal response mode. Exchanges over symmetric or balanced configurations can be set to specific mode using a frame design to deliver the command. HDLC offers three different modes of operation. These three modes of operations are:

- Normal Response Mode (NRM)
- Asynchronous Response Mode (ARM)
- Asynchronous Balanced Mode (ABM)

Normal Response Mode

This is the mode in which the primary station initiates transfers to the secondary station. The secondary station can only transmit a response when, and only when, it is instructed to do so by the primary station. In other words, the secondary station must receive explicit permission from the primary station to transfer a response. After receiving permission from the primary station, the secondary station initiates its transmission. This transmission from the secondary station to the primary station may be much more than just an acknowledgment of a frame. It may in fact be more than one information frame. Once the last frame is transmitted by the secondary station, it must wait once again from explicit permission to transfer anything, from the primary station. Normal Response Mode is only used within an unbalanced configuration.

Asynchronous Response Mode

In this mode, the primary station doesn't initiate transfers to the secondary station. In fact, the secondary station does not have to wait to receive explicit permission from the primary station to transfer any frames. The frames may be more than just acknowledgment frames. They may contain data, or control information regarding the status of the secondary station. This mode can reduce overhead on the link, as no frames need to be transferred in order to give the secondary station permission to initiate a transfer. However, some limitations do exist. Due to the fact that this mode is asynchronous, the secondary station must wait until it detects and idle channel before it can transfer any frames. This is when the ARM link is operating at half-duplex. If the ARM link is operating at full duplex, the secondary station can transmit at any time. In this mode, the primary station still retains responsibility for error recovery, link setup, and link disconnection.

Synchronous Balanced Mode

`This mode is used in case of combined stations. There is no need for permission on the part of any station in this mode. This is because combined stations do not require any sort of instructions to perform any task on the link.

Normal Response Mode is used most frequently in multi-point lines, where the primary station controls the link. Asynchronous Response Mode is better for point-to-point links, as it reduces overhead. Asynchronous Balanced Mode is not used widely today. The "asynchronous" in both ARM and ABM does not refer to the format of the

data on the link. It refers to the fact that any given station can transfer frames without explicit permission or instruction from any other station.

HDLC Non-Operational Modes

HDLC also defines three non-operational modes. These three non-operational modes are:

• Normal Disconnected Mode (NDM)
• Asynchronous Disconnected Mode (ADM)
• Initialization Mode (IM)

The two disconnected modes (NDM and ADM) differ from the operational modes in that the secondary station is logically disconnected from the link (note the secondary station is not physically disconnected from the link). The IM mode is different from the operations modes in that the secondary station's data link control program is in need of regeneration or it is in need of an exchange of parameters to be used in an operational mode.

HDLC Frame Structure

There are three different types of frames as shown in Fig. 3.4.4 and the size of different fields are shown Table 3.4.1.

Table 3.4.1 Size of different fields

Field Name	Size(in bits)
Flag Field(F)	8 bits
Address Field(A)	8 bits
Control Field(C)	8 or 16 bits
Information Field(I) OR Data	Variable; Not used in some frames
Frame Check Sequence(FCS)	16 or 32 bits
Closing Flag Field(F)	8 bits

The Flag field

Every frame on the link must begin and end with a flag sequence field (F). Stations attached to the data link must continually listen for a flag sequence. The flag sequence is an octet looking like 01111110. Flags are continuously transmitted on the link between frames to keep the link active. Two other bit sequences are used in HDLC as signals for the stations on the link. These two bit sequences are:

• Seven 1's, but less than 15 signal an abort signal. The stations on the link know there is a problem on the link.

• 15 or more 1's indicate that the channel is in an idle state.

The time between the transmissions of actual frames is called the interframe time fill. The interframe time fill is accomplished by transmitting continuous flags between frames. The flags may be in 8 bit multiples.

HDLC is a code-transparent protocol. It does not rely on a specific code for interpretation of line control. This means that if a bit at position N in an octet has a specific meaning, regardless of the other bits in the same octet. If an octet has a bit sequence of 01111110, but is not a flag field, HLDC uses a technique called bit-stuffing to differentiate this bit sequence from a flag field as we have discussed in the previous lesson.

At the receiving end, the receiving station inspects the incoming frame. If it detects 5 consecutive 1's it looks at the next bit. If it is a 0, it pulls it out. If it is a 1, it looks at the 8th bit. If the 8th bit is a 0, it knows an abort or idle signal has been sent. It then proceeds to inspect the following bits to determine appropriate action. This is the manner in which HDLC achieves code-transparency. HDLC is not concerned with any specific bit code inside the data stream. It is only concerned with keeping flags unique.

The Address field

The address field (A) identifies the primary or secondary stations involvement in the frame transmission or reception. Each station on the link has a unique address. In an unbalanced configuration, the A field in both commands and responses refer to the secondary station. In a balanced configuration, the command frame contains the destination station address and the response frame has the sending station's address.

The Control field

HDLC uses the control field (C) to determine how to control the communications process. This field contains the commands, responses and sequences numbers used to maintain the data flow accountability of the link, defines the functions of the frame and initiates the logic to control the movement of traffic between sending and receiving stations. There three control field formats:

• Information Transfer Format: The frame is used to transmit end-user data between two devices.

• Supervisory Format: The control field performs control functions such as acknowledgment of frames, requests for re-transmission, and requests for temporary suspension of frames being transmitted. Its use depends on the operational mode being used.

• Unnumbered Format: This control field format is also used for control purposes. It is used to perform link initialization, link disconnection and other link control functions.

The Poll/Final Bit (P/F)

The 5th bit position in the control field is called the poll/final bit, or P/F bit. It can only be recognized when it is set to 1. If it is set to 0, it is ignored. The poll/final bit is used to provide dialogue between the primary station and secondary station. The primary station uses P=1 to acquire a status response from the secondary station. The P bit signifies a poll. The secondary station responds to the P bit by transmitting a data or status frame to the primary station with the P/F bit set to F=1. The F bit can also be used to signal the end of a transmission from the secondary station under Normal Response Mode.

The Information field or Data field

This field is not always present in a HDLC frame. It is only present when the Information Transfer Format is being used in the control field. The information field contains the actually data the sender is transmitting to the receiver in an I-Frame and network management information in U-Frame.
The Frame check Sequence field

This field contains a 16-bit, or 32-bit cyclic redundancy check bits. It is used for error detection.

HDLC Commands and Responses

The set of commands and responses in HDLC is summarized in Table 3.4.2.

Information transfer format command and response (I-Frame)

The function of the information command and response is to transfer sequentially numbered frames, each containing an information field, across the data link.
Supervisory format command and responses (S-Frame)

Supervisory (S) commands and responses are used to perform numbered supervisory functions such as acknowledgment, polling, temporary suspension of information transfer, or error recovery. Frames with the S format control field cannot contain an information field. A primary station may use the S format command frame with the P bit set to 1 to request a response from a secondary station regarding its status. Supervisory Format commands and responses are as follows:

• Receive Ready (RR) is used by the primary or secondary station to indicate that it is ready to receive an information frame and/or acknowledge previously received frames.
• Receive Not Ready (RNR) is used to indicate that the primary or secondary station is not ready to receive any information frames or acknowledgments.
• Reject (REJ) is used to request the retransmission of frames.
• Selective Reject (SREJ) is used by a station to request retransmission of specific frames. An SREJ must be transmitted for each erroneous frame; each frame is treated as a separate error. Only one SREJ can remain outstanding on the link at any one time.

TABLE 3.4.2 HDLC Commands and Responses

Information Transfer	Information Transfer
Format Commands	Format Responses
I - Information	I - Information
Supervisory Format	Supervisory Format
Commands	Responses
RR - Receive ready	RR - Receive ready
RNR - Receive not ready	RNR - Receive not ready
REJ - Reject	REJ - Reject
SREJ - Selective reject	SREJ - Selective reject
Unnumbered Format	Unnumbered Format
Commands	Commands
SNRM - Set Normal Response Mode	UA - Unnumbered Acknowledgment
SARM - Set Asynchronous Response Mode	DM - Disconnected Mode
SABM - Set Asynchronous Balanced Mode	RIM - Request Initialization Mode
DISC - Disconnect	RD - Request Disconnect
SNRME - Set Normal Response Mode Extended	UI - Unnumbered Information
SARME - Set Asynchronous Response Mode Extended	XID - Exchange Identification
SABME - Set Asynchronous Balanced Mode Extended	FRMR - Frame Reject
SIM - Set Initialization Mode	TEST - Test
UP - Unnumbered Poll	
UI - Unnumbered Information	

XID - Exchange identification
RSET - Reset
TEST - Test

Unnumbered Format Commands and responses (U-Frame)

The unnumbered format commands and responses are used to extend the number of data link control functions. The unnumbered format frames have 5 modifier bits, which allow for up to 32 additional commands and 32 additional response functions. Below, 13 command functions, and 8 response functions are described.

• Set Normal Response Mode (SNRM) places the secondary station into NRM. NRM does not allow the secondary station to send any unsolicited frames. Hence the primary station has control of the link.

• Set Asynchronous Response Mode (SARM) allows a secondary station to transmit frames without a poll from the primary station.

• Set Asynchronous Balanced Mode (SABM) sets the operational mode of the link to ABM.

• Disconnect (DISC) places the secondary station in to a disconnected mode.

• Set Normal Response Mode Extended (SNRME) increases the size of the control field to 2 octets instead of one in NRM. This is used for extended sequencing. The same applies for SARME and SABME.

• Set Initialization Mode (SIM) is used to cause the secondary station to initiate a station-specific procedure(s) to initialize its data link level control functions.

• Unnumbered Poll (UP) polls a station without regard to sequencing or acknowledgment.

• Unnumbered Information (UI) is used to send information to a secondary station.

• Exchange Identification (XID) is used to cause the secondary station to identify itself and provide the primary station identifications characteristics of itself.

• Reset (RSET) is used to reset the receive state variable in the addressed station.

• Test (TEST) is used to cause the addressed secondary station to respond with a TEST response at the first response opportunity. It performs a basic test of the data link control.

• Unnumbered Acknowledgment (UA) is used by the secondary station to acknowledge the receipt and acceptance of an SNRM, SARM, SABM, SNRME, SARME, SABME, RSET, SIM, or DISC commands.

• Disconnected Mode (DM) is transmitted from a secondary station to indicate it is in disconnected mode(non-operational mode.)

• Request Initialization Mode (RIM) is a request from a secondary station for initialization to a primary station. Once the secondary station sends RIM, it can only respond to SIM, DSIC, TEST or XID commands.

• Request Disconnect (RD) is sent by the secondary station to inform the primary station that it wishes to disconnect from the link and go into a non-operational mode(NDM or ADM.

• Frame Reject (FRMR) is used by the secondary station in an operation mode to report that a condition has occurred in transmission of a frame and retransmission of the frame will not correct the condition.

HDLC Subsets

Many other data link protocols have been derived from HDLC. However, some of them reach beyond the scope of HDLC. Two other popular offsets of HDLC are Synchronous Data Link Control (SDLC), and Link Access Protocol, Balanced (LAP-B). SDLC is used and developed by IBM. It is used in a variety of terminal to computer applications. It is also a part of IBM's SNA communication architecture. LAP-B was developed by the ITU-T. It is derived mainly from the asynchronous response mode (ARM) of HDLC. It is commonly used for attaching devices to packet-switched networks.

• Combined Station: A combined station is a combination of a primary and secondary station. On the link, all combined stations are able to send and receive commands and responses without any permission from any other stations on the link.

Specific Instructional Objectives

At the end of this lesson, the students will become familiar with the following concepts:

• Explain the operation of IEEE 802 LANs
o 802.4 – Token bus-based
o 802.5 – Token ring-based
• Compare performance of the three LANs

2.7.Introduction

In the preceding lesson we have mentioned that for the fulfillment of different goals, the IEEE 802 committee came up with a bunch of LAN standards collectively known as LANs as shown in Fig. 5.4.1. We have already discussed CSMA/CD-based LAN proposed by the IEEE 802.3 subcommittee, commonly known as Ethernet. In this lesson we shall discuss Token bus, Token Ring based LANs proposed by the IEEE 802.4 and IEEE 8.2.5 subcommittees.
Specific Instructional Objectives

At the end of this lesson, the students will become familiar with the following concepts:

• Explain the operation of IEEE 802 LANs
o 802.4 – Token bus-based
o 802.5 – Token ring-based
• Compare performance of the three LANs

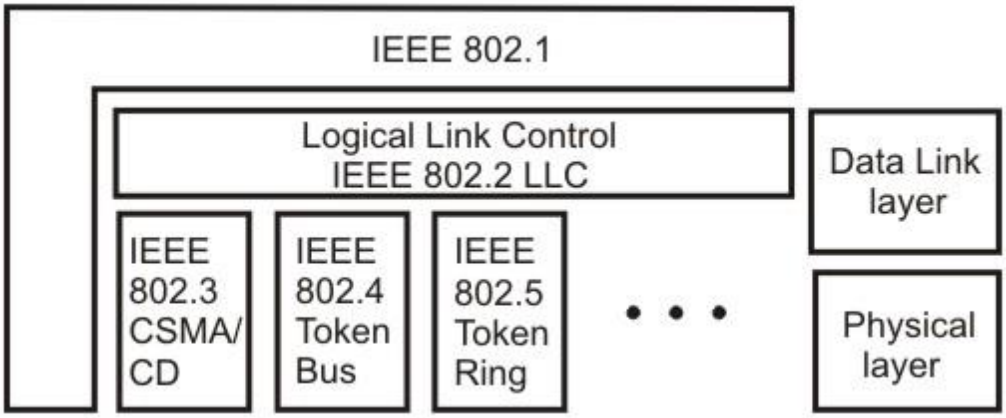


Figure 5.4.1 IEEE 802 Legacy LANs

2.8. Token Ring (IEEE 802.5)

Token Ring: A Brief History

Originally, IBM developed Token Ring network in the 1970s. It is still IBM's primary local-area network (LAN) technology. The related IEEE 802.5 specification is almost identical to and completely compatible with IBM's Token Ring network. In fact, the IEEE 802.5 specification was modeled after IBM Token Ring, and on the same lines. The term *Token Ring* is generally used to refer to both IBM's Token Ring network and IEEE 802.5 networks.

Introduction

Before going into the details of the Token Ring protocol, let's first discuss the motivation behind it. As already discussed, the medium access mechanism used by Ethernet (CSMA/CD) may results in collision. Nodes attempt to a number of times before they can actually transmit, and even when they start transmitting there are chances to encounter collisions and entire transmission need to be repeated. And all this become worse one the traffic is heavy i.e. all nodes have some data to transmit. Apart from this there is no way to predict either the occurrence of collision or delays produced by multiple stations attempting to capture the link at the same time. So all these problems with the Ethernet gives way to an alternate LAN technology, Token Ring.

Token Ring and IEEE802.5 are based on token passing MAC protocol with ring topology. They resolve the uncertainty by giving each station a turn on by one. Each node takes turns sending the data; each station may transmit data during its turn. The technique that coordinates this turn mechanism is called Token passing; as a Token is passed in the network and the station that gets the token can only transmit. As one node transmits at a time, there is no chance of collision. We shall discuss the detailed operation in next section.

Stations are connected by point-to-point links using repeaters. Mainly these links are of shielded twisted-pair cables. The repeaters function in two basic modes: Listen mode, Transmit mode. A disadvantage of this topology is that it is vulnerable to link or station failure. But a few measures can be taken to take care of it.

Differences between Token Ring and IEEE 802.5

Both of these networks are basically compatible, although the specifications differ in some ways.

• IEEE 802.5 does not specify a topology, although virtually all IEEE 802.5 implementations are based on the star topology. While IBM's Token Ring network explicitly specifies a star, with all end stations attached to a device called a Multi-Station Access Unit (MSAU).

• IEEE 802.5 does not specify a media type, although IBM Token Ring networks use twisted-pair wire.

• There are few differences in routing information field size of the two.

Token Ring Operation

Token-passing networks move a small frame, called a *token*, around the network. Possession of the token grants the right to transmit. If a node receiving the token has no information to send, it passes the token to the next end station. Each station can hold the token for a maximum period of time.

If a station possessing the token does have information to transmit, it seizes the token, alters 1 bit of the token (which turns the token into a start-of-frame sequence), appends the information that it wants to transmit, and sends this information to the next station on the ring. While the information frame is circling the ring, no token is on the network (unless the ring supports early token release), which means that other stations wanting to transmit must wait. Therefore, *collisions cannot occur in Token Ring networks*. If *early token release* is supported, a new token can be released immediately after a frame transmission is complete.

The information frame circulates around the ring until it reaches the intended destination station, which copies the information for further processing. The information frame makes a round trip and is finally removed when it reaches the sending station. The sending station can check the returning frame to see whether the frame was seen and subsequently copied by the destination station in error-free form. Then the sending station inserts a new free token on the ring, if it has finished transmission of its packets.

Unlike CSMA/CD networks (such as Ethernet), token-passing networks are *deterministic*, which means that it is possible to calculate the maximum time that will pass before any end station will

be capable of transmitting. Token Ring networks are ideal for applications in which delay must be predictable and robust network operation is important.

Priority System

Token Ring networks use a sophisticated priority system that permits certain user-designated, high-priority stations to use the network more frequently. Token Ring frames have two fields that control priority: *the priority field* and the *reservation field.*

Only stations with a priority equal to or higher than the priority value contained in a token can seize that token. After the token is seized and changed to an information frame, only stations with a priority value higher than that of the transmitting station can reserve the token for the next pass around the network. When the next token is generated, it includes the higher priority of the reserving station. Stations that raise a token's priority level must reinstate the previous priority after their transmission is complete.

Ring Maintenance

There are two error conditions that could cause the token ring to break down. One is the *lost token* in which case there is no token the ring, the other is the *busy token* that circulates endlessly. To overcome these problems, the IEEE 802 standard specifies that one of the stations be designated as 'active monitor'. The monitor detects the lost condition using a timer by *time-out* mechanism and recovers by using a new free token. To detect a circulating busy token, the monitor sets a 'monitor bit' to one on any passing busy token. If it detects a busy token with the monitor bit already set, it implies that the sending station has failed to remove its packet and recovers by changing the busy token to a free token. Other stations on the ring have the role of passive monitor. The primary job of these stations is to detect failure of the active monitor and assume the role of active monitor. A contention-resolution is used to determine which station to take over.

.6 Physical Layer

The Token Ring uses shielded twisted pair of wire to establish point-point links between the adjacent stations. The baseband signaling uses differential Manchester encoding. To overcome the problem of cable break or network failure, which brings the entire network down, one suggested technique, is to use *wiring concentrator* as shown in Fig. 5.4.2.

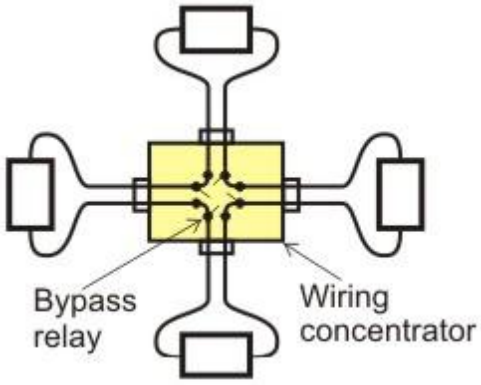


Figure 5.4.2 Star Connected Ring topology

It imposes the reliability in an elegant manner. Although logically the network remains as a ring, physically each station is connected to the *wire center* with two twisted pairs for 2-way communication. Inside the wire center, *bypass relays* are used to isolate a broken wire or a faulty station. This Topology is known as *Star-Connected Ring***.**

7 Frame Format

Token Ring and IEEE 802.5 support two basic frame types: tokens and data/command frames. Tokens are 3 bytes in length and consist of a start delimiter, an access control byte, and an end delimiter. Data/command frames vary in size, depending on the size of the Information field. Data frames carry information for upper-layer protocols, while command frames contain control information and have no data for upper-layer protocols.

Token Frame Fields

Start Delimiter	Access Control	Ending delimiter

Token Frame contains three fields, each of which is 1 byte in length:

• **Start delimiter (1 byte)**: Alerts each station of the arrival of a token (or data/command frame). This field includes signals that distinguish the byte from the rest of the frame by violating the encoding scheme used elsewhere in the frame.

• **Access-control (1 byte)**: Contains the Priority field (the most significant 3 bits) and the Reservation field (the least significant 3 bits), as well as a token bit (used to differentiate a token from a data/command frame) and a monitor bit (used by the active monitor to determine whether a frame is circling the ring endlessly).

• **End delimiter (1 byte)**: Signals the end of the token or data/command frame. This field also contains bits to indicate a damaged frame and identify the frame that is the last in a logical sequence.

Data/Command Frame Fields

Start Delimiter	Access Control	Frame Control	Destination address	Source address	Data	Frame check sequence	End Delimiter	Frame Status

Data/command frames have the same three fields as Token Frames, plus several others. The Data/command frame fields are described below:

• **Frame-control byte (1 byte)**—Indicates whether the frame contains data or control information. In control frames, this byte specifies the type of control information.

• **Destination and source addresses (2-6 bytes)**—Consists of two 6-byte address fields that identify the destination and source station addresses.

• **Data (up to 4500 bytes)**—Indicates that the length of field is limited by the ring token holding time, which defines the maximum time a station can hold the token.

• **Frame-check sequence (FCS- 4 byte)**—Is filed by the source station with a calculated value dependent on the frame contents. The destination station recalculates the value to determine whether the frame was damaged in transit. If so, the frame is discarded.

• **Frame Status (1 byte)**—This is the terminating field of a command/data frame. The Frame Status field includes the address-recognized indicator and frame-copied indicator.

2.9 Token Bus (IEEE 802.4)

1. Token BUS: A Brief History

Although Ethernet was widely used in the offices, but people interested in factory automation did not like it because of the probabilistic MAC layer protocol. They wanted a protocol which can support priorities and has predictable delay. These people liked the conceptual idea of Token Ring network but did not like its physical implementation as a break in the ring cable could bring the whole network down and ring is a poor fit to their linear assembly lines. Thus a new standard, known as Token bus, was developed, having the robustness of the Bus topology, but the known worst-case behavior of a ring.

Here stations are logically connected as a ring but physically on a Bus and follows the collision-free token passing medium access control protocol. So the motivation behind token bus protocol can be summarized as:

- The probabilistic nature of CSMA/ CD leads to uncertainty about the delivery time; which created the need for a different protocol
- The token ring, on the hand, is very vulnerable to failure.
- Token bus provides deterministic delivery time, which is necessary for real time traffic.
- Token bus is also less vulnerable compared to token ring.

2. Functions of a Token Bus

It is the technique in which the station on bus or tree forms a logical ring, that is the stations are assigned positions in an ordered sequence, with the last number of the sequence followed by the first one as shown in Fig. 5.4.3. Each station knows the identity of the station following it and preceding it.

Figure 5.4.3 Token Bus topology

A control packet known as a *Token* regulates the right to access. When a station receives the token, it is granted control to the media for a specified time, during which it may transmit one or more packets and may poll stations and receive responses when the station is done, or if its time has expired then it passes token to next station in logical sequence. Hence, steady phase consists of alternate phases of token passing and data transfer.

The MAC sublayer consists of four major functions: the interface machine (IFM), the access control machine (ACM), the receiver machine (RxM) and the transmit machine (TxM).

IFM interfaces with the LLC sublayer. The LLC sublayer frames are passed on to the ACM by the IFM and if the received frame is also an LLC type, it is passed from RxM component to the LLC sublayer. IFM also provides quality of service.

The **ACM** is the heart of the system. It determines when to place a frame on the bus, and responsible for the maintenance of the logical ring including the *error detection* and *fault recovery*.

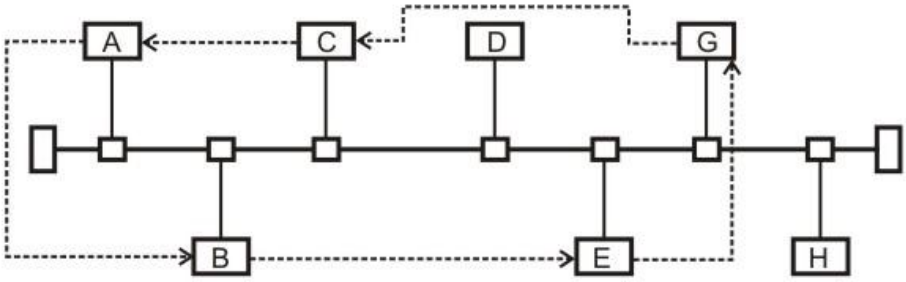


Figure 5.4.3 Token Bus topology

A control packet known as a *Token* regulates the right to access. When a station receives the token, it is granted control to the media for a specified time, during which it may transmit one or more packets and may poll stations and receive responses when the station is done, or if its time has expired then it passes token to next station in logical sequence. Hence, steady phase consists of alternate phases of token passing and data transfer.

The MAC sublayer consists of four major functions: the interface machine (IFM), the access control machine (ACM), the receiver machine (RxM) and the transmit machine (TxM).

IFM interfaces with the LLC sublayer. The LLC sublayer frames are passed on to the ACM by the IFM and if the received frame is also an LLC type, it is passed from RxM component to the LLC sublayer. IFM also provides quality of service.

The **ACM** is the heart of the system. It determines when to place a frame on the bus, and responsible for the maintenance of the logical ring including the *error detection* and *fault recovery*. It also cooperates with other stations ACM's to control the access to the shared bus, controls the admission of new stations and attempts recovery from faults and failures.

The responsibility of a **TxM** is to transmit frame to physical layer. It accepts the frame from the ACM and builds a MAC protocol data unit (PDU) as per the format.

The **RxM** accepts data from the physical layer and identifies a full frame by detecting the SD and ED (start and end delimiter). It also checks the FCS field to validate an error-free transmission.

3. Frame Form

The frame format of the Token Bus is shown in Fig. 5.4.4. Most of the fields are same as Token Ring. So, we shall just look at the Frame Control Field in Table 5.4.1

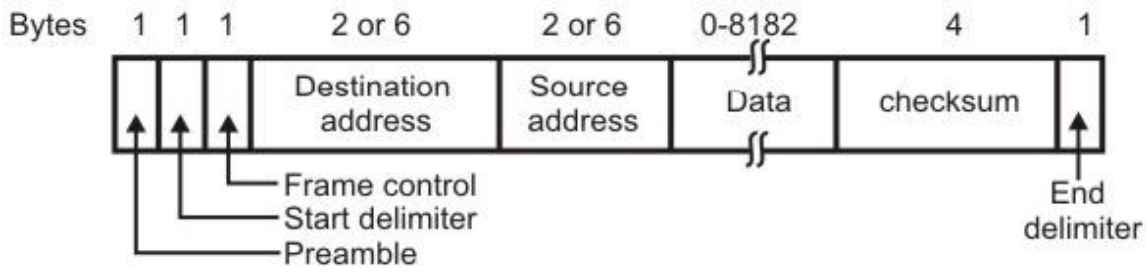


Figure 5.4.4 Token Bus frame format

Table 5.4.1 The Frame Control Field

Frame Control	**Name**	**Use**
0000 0000	Claim-Token	Ring Initialization
0000 0001	Solicit-successor -1	Addition to the Ring
0000 0010	Solicit-successor -2	Addition to the Ring
0000 0011	Who-follows	Recovery from lost token
0000 0100	Resolve Contention	Multiple station to join the Ring
0000 1000	Token	Pass the Token
0000 1100	Set-Successor	Deletion from the ring

4. Logical ring maintenance

The MAC performs the following functions as part of its maintenance role of the ring.

Addition to the Ring: Non-participating stations must periodically be granted the opportunity to insert themselves into the ring. Each node in the ring periodically grants an opportunity for new nodes to enter the ring while holding the token. The node issues a solicit–successor–1 packet, inviting nodes with an address between itself and the next node in logical sequence to request entrance. The transmitting node then waits for a period of time equal to one response window or slot time (twice the end-to-end propagation delay of the medium). If there is no request, the token holder sets its successor node to be the requesting node and transmits the token to it; the requester sets the linkages accordingly and proceeds.

If more than one node requests, to enter the ring, the token holder will detect a garbled transmission. The conflict is resolved by *addressed based contention scheme;* the token holder transmits a resolved contention packet and waits for four response windows. Each requester can transmit in one of these windows, based on the first two bits of its address.

If requester hears anything before its windows comes up, it refrains from requesting entrance. If a token holder receives a valid response, then it can proceed, otherwise it tries again and only those nodes that request the first time are allowed to request this time, based on the second pair of bits in their address. This process continues until a valid request is received or no request is received, or a maximum retry count is reached. In latter cases, the token holder passes the token to logical successor in the ring.

Deletion from Ring: A station can voluntarily remove itself from the ring by splicing together its predecessor and successor. The node which wants to be deleted from the ring waits until token comes to it, then it sends a set successor packet to its predecessor, instructing it to splice to its successor.

Fault Management: Errors like duplicate address or broken ring can occur. A suitable management scheme should be implemented for smooth functioning. It is done by the token-holder first, while holding the token, node may hear a packet, indicating that another node has the token. In this case, it immediately drops the token by reverting to listener mode, and the number of token holders drops immediately from one to zero. Upon completion of its turn, it immediately issues a data or token packet. The sequence of steps are as follows:

i. After sending the token, the token issuer will listen for one slot time to make sure that its predecessor is active.

ii. If the issuer does not hear a valid packet, it reissues the token to the same successor one more time.

iii. After two failures, the issuer assumes that its successor has failed and issues a "who-follows" packet, asking for the identity of the node that follows the failed node. The issuer should get back a set successor packet from the second node down the time. If so, the issuer adjusts its linkage and issues a token (back to step i).

iv. If the issuing node gets a response to its "who-follows" packet, it tries again.

v. If the "who-follows" tactic fails, the node issues a solicit-successor-2 packet with full address range (i.e. every node is invited to respond). If this packet works then the ring is established and procedure continues.

vi. If two attempts in step (v) fail, it assumes that a catastrophe has happened; perhaps the node receiver has failed. In any case, the node ceases the activity and listen the bus.

Ring Initialization: Ring is to be initialized by starting the token passing. This is necessary when the network is being setup or when ring is broken down. Some decentralized algorithms should take care of, who starts first, who starts second, etc. it occurs when one or more stations detects a lack of bus activity lasting longer than a specific time. The token may get lost. This can occur on a number of occasions. For example, when network has been just powered up, or a token holding station fails. Once its time out expires, a node will issue a claim token packet. Contending clients are removed in a similar fashion to the response window process.

4. Relative comparison of the three standards

A comparison of the three standards for different functions is shown in Table 5.4.2 and results of the analysis of the performance of the three standards are summarized below:

• The CSMA/CD protocol shows strong dependence on the parameter 'a', which is the ratio of the propagation time to the transmission time. It offers shortest delay under light load and it is most sensitive under heavy load conditions.

• Token ring is least sensitive to different load conditions and different packet sizes.

• Token bus is highly efficient under light load conditions.

Table 5.4.2 Comparison of the three standards

Function	CSMA/CD	**Token bus**	**Token ring**
Access determination	Contention	Token	Token
Packet length restriction	64 bytes (Greater than 2.Tprop)	None	None
Priority	Not supported	Supported	Supported
Sensitivity to work load	Most sensitive	Sensitive	Least sensitive
Principle advantage	Simplicity, wide installed base	Regulated/fair access	Regulated/fair access
Principle disadvantage	Nondeterministic delay	Complexity	Complexity

2.10. Fast Ethernet

The 802.u or the fast Ethernet, as it is commonly known, was approved by the IEEE 802 Committee in June 1995. It may not be considered as a new standard but an addendum to the existing 802.3 standard. The fast Ethernet uses the same frame format, same CSMA/CD protocol and same interface as the 802.3, but uses a data transfer rate of 100 Mb/s instead of 10 Mb/s. However, fast Ethernet is based entirely on 10-Base-T, because of its advantages (Although technically 10-BASE-5 or 10-BASE-2 can be used with shorter segment length).

Fortunately, the Ethernet is designed in such a way that the speed can be increased if collision domain is decreased. The only two changes made in the MAC layer are the data

rate and the collision domain. The data rate is increased by a factor of 10 and collision domain is decreased by a factor of 10. To increase the data rate without changing the minimum size of the frame (576 bits or 76 bytes in IEEE 802.3), it is necessary to decrease the round-trip delay time. With the speed of 100Mbps the round-trip time reduce to 5.76 microseconds (576 bits/100 Mbps; which was 57.6 microsecond for 10Mbps Normal Ethernet). This means that the collision domain is decreased 10 fold from 2500 meters (in IEEE802.3) to 250 meters (fast Ethernet).

IEEE has designed two categories of Fast Ethernet: 100Base-X and 100Base-T4. 100Base-X uses two-wire interface between a hub and a station while 100Base-T4 uses four-wire interface. 100-Base-X itself is divided into two: 100Base-TX and 100base-FX as shown in Fig. 5.6.2.

100 BASE-T4:

This option is designed to avoid overwriting. It is used for half-duplex communication using four wire-pairs of the existing category 3 UTP cables, which are already available for telephone services in homes/offices. Two of four pairs are bi-directional; other two are unidirectional. This means that there are 3 pairs to be used for carrying data, in each direction (2 bi-directional and 1 uni-directional) as shown in Fig. 5.6.3. Because 100Mbps data cannot be handled by voice-grade UTP, this specification splits the 100 Mbps flow into three 33.66 Mbps flows.

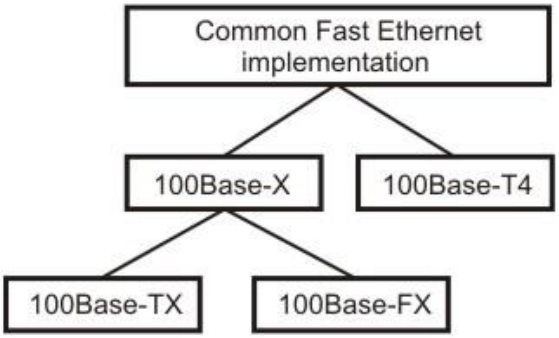


Figure 5.6.2 Fast Ethernet implementations

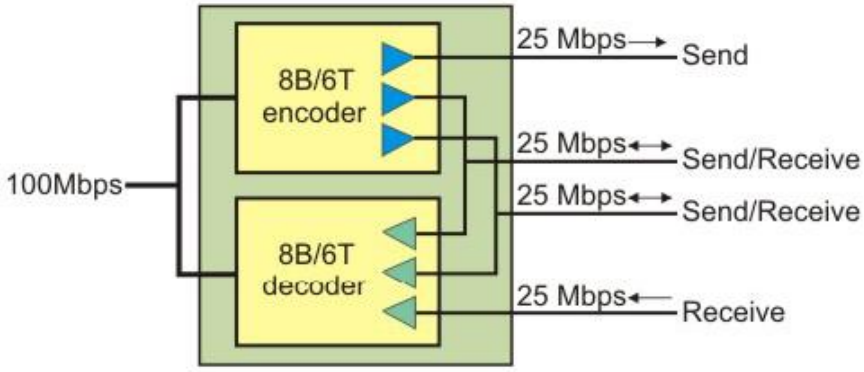


Figure 5.6.3 100Base-T4 implementation

100 BASE TX:

This option uses two pairs of category 5 UTP or two shielded twisted-pair (STP) cable to connect a station to hub as shown in Fig. 5.6.4. One pair is used to carry frames from the hub to the station and other to carry frames from station to hub. It uses 4B/5B encoding to handle 100 Mbps using NRZ-I signaling. The distance between station and hub should be less than 100 meters.

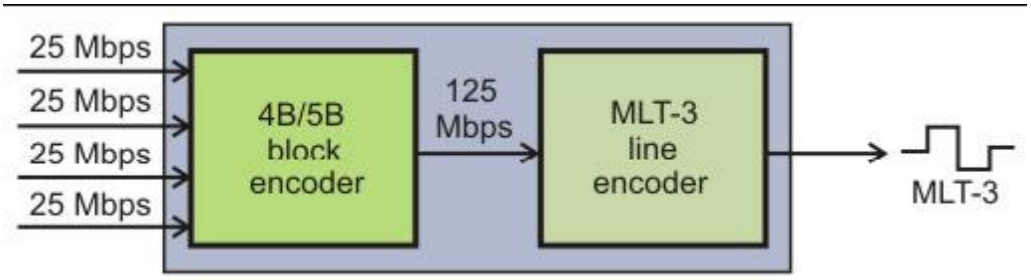


Figure 5.6.4 100Base-TX implementation

100 BASE FX:

This option uses two Fiber optic cables, one carry frames from station to hub and other from hub to station as shown in Fig. 5.6.5. The encoding is using 4B/5B and it uses NRZ-I signaling. The distance between station and hub should be less than 2000 meters.

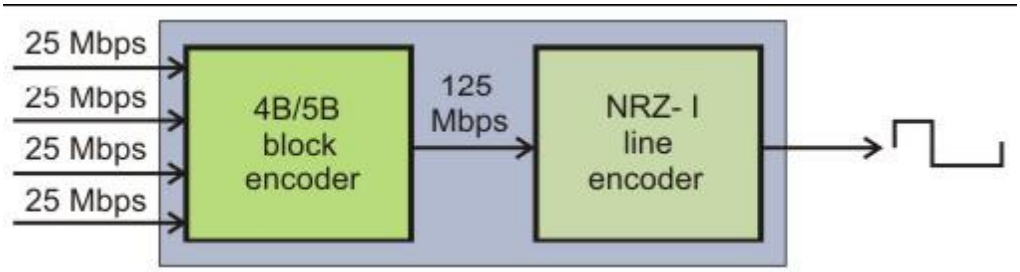


Figure 5.6.5 100Base-FX implementation

2.11. Gigabit Ethernet

1. Brief History and the IEEE 802.3z Task Force

As applications increased, the demand on the network, newer, high-speed protocols such as FDDI and ATM became available. However, in the last couple of years, Fast Ethernet has become the backbone of choice because it's simplicity and its reliance on Ethernet. The primary goal of Gigabit Ethernet is to build on that topology and knowledge base to build a higher-speed protocol without forcing customers to throw away existing networking equipment.

In March 1996, the IEEE 802.3 committee approved the 802.3z Gigabit Ethernet Standardization project. At that time as many as 54 companies expressed there intent to participate in the standardization project. The Gigabit Ethernet Alliance was formed in May 1996 by 11 companies. The Alliance represents a multi-vendor effort to provide open and inter-operable Gigabit Ethernet products. The objectives of the alliance are:

• Supporting extension of existing Ethernet and Fast Ethernet technology in response to demand for higher network bandwidth.
• Developing technical proposals for the inclusion in the standard
• Establishment of inter-operability test procedures and processes

2. Similarities and advances over Ethernet (IEEE 802.3)

As its name implies, Gigabit Ethernet - officially known as 802.3z - is the 1 Gb/s extension of the 802.3 standard already defined for 10 and 100 Mb/s service. Gigabit Ethernet builds on top of the Ethernet protocol, but increases speed tenfold over Fast Ethernet to 1000 Mbps, or 1 gigabit per second (Gbps). It retains the Carrier Sense Multiple Access/ Collision Detection (CSMA/CD) as the access method. It supports full duplex as well as half duplex modes of operation. Initially, single-mode and multi mode fiber and short-haul coaxial cable were supported. Standards for twisted pair cables were subsequently added. The standard uses physical signaling technology used in Fiber Channel to support Gigabit rates over optical fibers. Since Gigabit Ethernet significantly leverages on Ethernet, customers will be able to leverage their existing knowledge base to manage and maintain gigabit networks. Initially, Gigabit Ethernet was expected to be used as a backbone system in existing networks. It can be used to aggregate traffic between clients and "server farms", and for connecting Fast Ethernet switches. It can also be used for connecting workstations and servers for high-bandwidth applications such as medical imaging or CAD. But, gigabit Ethernet is not simply a straight Ethernet running at 1 Gb/s. In fact, the ways it differs from its predecessors may be more important than its similarities. Some of the important differences are highlighted below.

(i) The cabling requirement of gigabit Ethernet is very different. The technology is based on fiber optic cable. Multi-mode fiber is able to transmit at gigabit rate to at least 580 meters and with single-mode runs exceeding 3 km. Fiber optic cabling is costly. In order to reduce the cost of cabling, the 802.3z working group also proposed the use of twisted-pair or cable or coaxial cable for distances up to 30 meters.

(ii) Gigabit Ethernet also relies on a modified MAC layer. At gigabit speed, two stations 200 meters apart will not detect a collision, when both simultaneously send 64-byte frames. This inability to detect collision leads to network instability. A mechanism known as *carrier extension* has been proposed for frames shorter than 512 bytes. The number of repeater hops is also restricted to only one in place of two for 100 Base-T.

(iii) Flow Control is a major concern in gigabit Ethernet because of buffer overflow and junked frames in heavily loaded condition. The solution proposed by IEEE subcommittee is the 802.3x. The X-on/X-off protocol works over any full-duplex Ethernet, fast Ethernet or gigabit Ethernet link. When a switch buffer is close to capacity, the receiving device signals the sending station and tells it to stop transmitting until the buffer becomes empty.

(iv) Finally, one important feature, which Ethernet technology lacks, is the Quality of Service (QoS). The gigabit Ethernet is a connectionless technology that transmits variable length frames. As such, it simply cannot guarantee that the real-time packets get the preferential treatment they require. The IEEE subcommittee developed two specifications

that will help Ethernet provide the required QoS. 802.lq tags traffic for VLANs and for prioritization. 802.lp is a signaling scheme that lets end station request priority and allows switches to pass these requests along the path.

The gigabit Ethernet comes into its own as an internetworking switch link (ISL) that aggregates 10-and100-Mb/s feeds from the desktops and servers. Presently, gigabit Ethernet is already matured with a large installation base as a backbone network technology.

3. Gigabit Ethernet Protocol Architecture

In order to accelerate speeds from 100 Mbps Fast Ethernet up to 1 Gbps, several changes were required to be made to the physical interface. It was decided that Gigabit Ethernet will look identical to Ethernet from the data link layer upward. The challenges involved in accelerating to 1 Gbps have been resolved by merging two technologies together: IEEE 802.3 Ethernet and ANSI X3T11 FiberChannel as shown in Fig. 5.6.6.

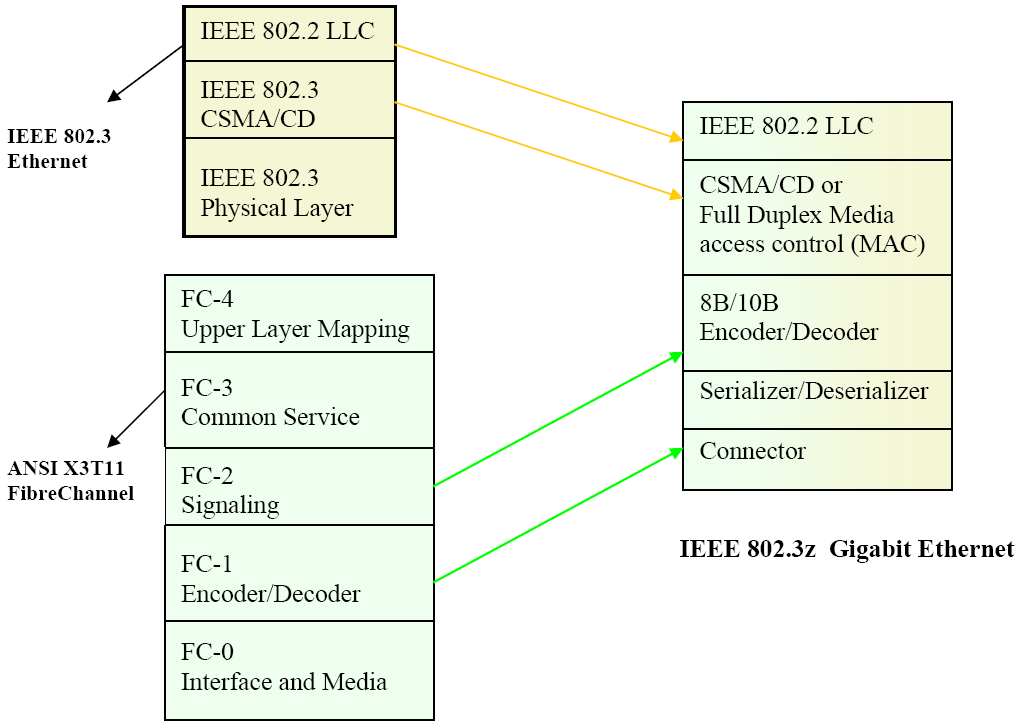


Figure 5.6.6 Gigabit Ethernet Architecture -1

4. GMII (Gigabit Media Independent Interface)

The various layers of the Gigabit Ethernet protocol architecture are shown in Fig. 5.6.7. The GMII is the interface between the MAC layer and the Physical layer. It allows any physical layer to be used with the MAC layer. It is an extension of the MII (Media

Independent Interface) used in Fast Ethernet. It uses the same management interface as MII. It supports 10, 100 and 1000 Mbps data rates. It provides separate 8-bit wide receive and transmit data paths, so it can support both full duplex as well as half duplex operation.

The GMII provides 2 media status signals: one indicates presence of the carrier, and the other indicates absence of collision. The Reconciliation Sublayer (RS) maps these signals to Physical Signaling (PLS) primitives understood by the existing MAC sublayer. With the GMII, it is possible to connect various media types such as shielded and unshielded twisted pair, and single-mode and multi mode optical fiber, while using the same MAC controller.

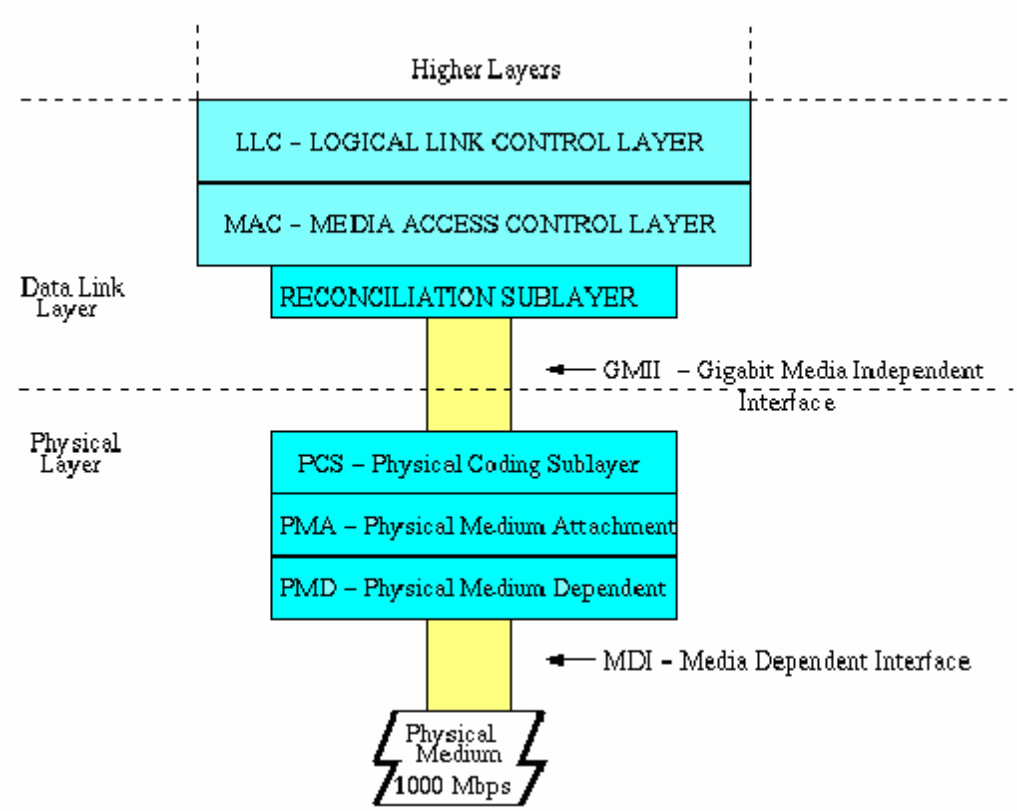


Figure 5.6.7 Gigabit Ethernet Architecture-2

• **PCS (Physical Coding Sublayer)**

This is the GMII sublayer, which provides a uniform interface to the Reconciliation layer for all physical media. It uses 8B/10B coding like Fiber Channel. In this type of coding, groups of 8 bits are represented by 10 bit "code groups". Some code groups represent 8-bit data symbols. Others are control symbols. The extension symbols used in Carrier Extension are an example of control symbols. Carrier Sense and Collision Detect indications are generated by this sublayer. It also manages the auto-negotiation process by which the NIC (Network Interface) communicates with the network to determine the network speed (10,100 or 1000 Mbps) and mode of operation (half-duplex or full-duplex).

• **PMA (Physical Medium Attachment)**

This sublayer provides a medium-independent means for the PCS to support various serial bit-oriented physical media. This layer serializes code groups for transmission and deserializes bits received from the medium into code groups.

• **PMD (Physical Medium Dependent)**

This sublayer maps the physical medium to the PCS. This layer defines the physical layer signalling used for various media. The **MDI (Medium Dependent Interface)**, which is a part of PMD, is the actual physical layer interface. This layer defines the actual physical attachment, such as connectors, for different media types divided into three sub layers: PCS, PMA and PMD.

5. Media Access Control Layer

Gigabit Ethernet has been designed to adhere to the standard Ethernet frame format. This setup maintains compatibility with the installed base of Ethernet and Fast Ethernet products, requiring no frame translation. Gigabit Ethernet maintains the minimum and maximum frame sizes of Ethernet. Since, Gigabit Ethernet is 10 times faster than Fast Ethernet, to maintain the same slot size, maximum cable length would have to be reduced to about 10 meters, which is not very useful. Instead, Gigabit Ethernet uses a bigger slot size of 512 bytes (In Ethernet, the slot size is 64 bytes, the minimum frame length). To maintain compatibility with Ethernet, the minimum frame size is not increased, but the "carrier event" is extended. If the frame is shorter than 512 bytes, then it is padded with extension symbols. These are special symbols, which cannot occur in the payload. This process is called *Carrier Extension*

• **Carrier Extension**

Gigabit Ethernet should be inter-operable with existing 802.3 networks. Carrier Extension is a way of maintaining 802.3 minimum and maximum frame sizes with meaningful cabling distances.

Figure 5.6.8 Ethernet Frame Format With Carrier Extension

For carrier extended frames, the non-data extension symbols are included in the "collision window", that is, the entire extended frame is considered for collision and dropped. However, the Frame Check Sequence (FCS) is calculated only on the original (without extension symbols) frame. The extension symbols are removed before the FCS is checked by the receiver. So the LLC (Logical Link Control) layer is not even aware of the carrier extension. Figure 5.6.8 shows the Ethernet frame format when Carrier Extension is used.

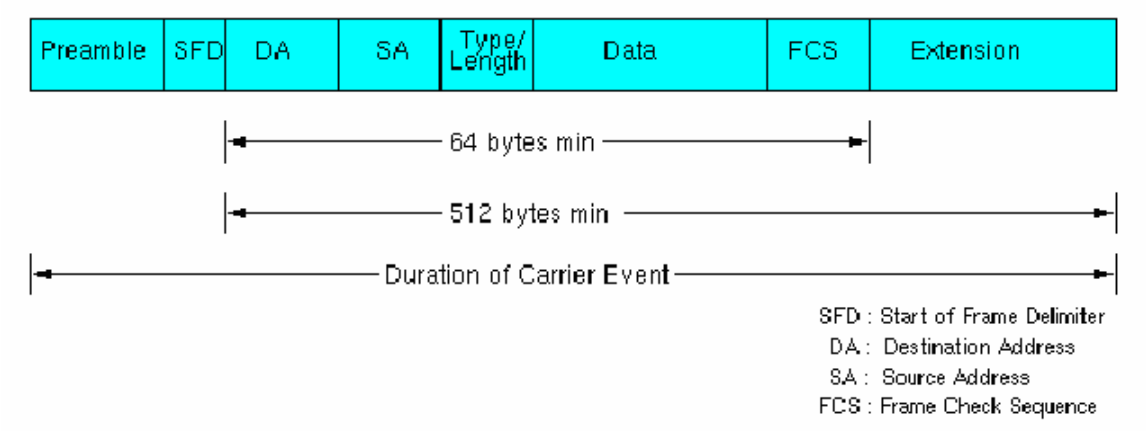


Figure 5.6.8 Ethernet Frame Format With Carrier Extension

For carrier extended frames, the non-data extension symbols are included in the "collision window", that is, the entire extended frame is considered for collision and dropped. However, the Frame Check Sequence (FCS) is calculated only on the original (without extension symbols) frame. The extension symbols are removed before the FCS is checked by the receiver. So the LLC (Logical Link Control) layer is not even aware of the carrier extension. Figure 5.6.8 shows the Ethernet frame format when Carrier Extension is used.

•Packet Bursting

Carrier Extension is a simple solution, but it wastes bandwidth. Up to 448 padding bytes may be sent for small packets. This results in lower throughput. In fact, for a large number of small packets, the throughput is only marginally better than Fast Ethernet.

Packet Bursting is an extension of Carrier Extension. Packet Bursting is "Carrier Extension plus a burst of packets". When a station has a number of packets to transmit, the first packet is padded to the slot time if necessary using carrier extension. Subsequent packets are transmitted back to back, with the minimum Inter-packet gap (IPG) until a burst timer (of 1500 bytes) expires. Packet Bursting substantially increases the throughput. Figure 5.6.9 shows how Packet Bursting works.

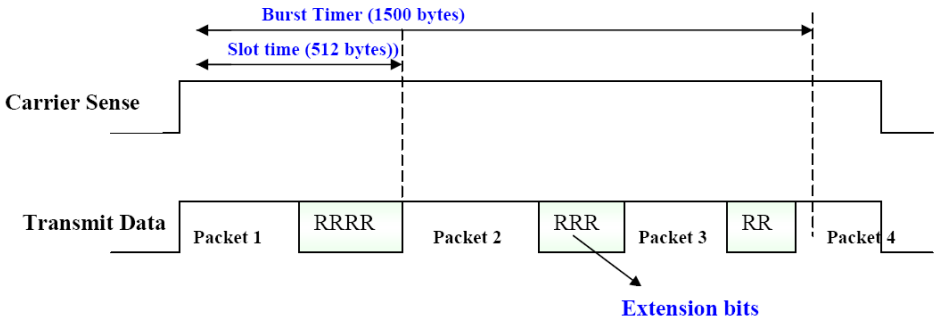


Figure 5.6.9 Packet Bursting

GBIC: Gigabit Ethernet Interface Carrier allows network managers to configure each port on a port-by-port basis, including long-haul (LH) to support a distance of 5-10 Km using SMF as shown in Fig. 5.6.10.

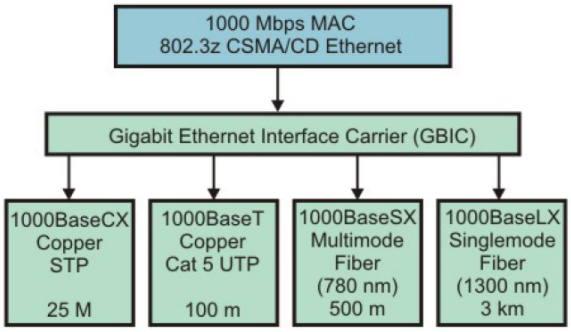


Figure 5.6.10 GBIC architecture

Migration to Gigabit Ethernet

Possible migration approaches to Gigabit Ethernet network from existing Fast Ethernet or Ethernet network is given below:

- Upgrading Switch-to-Switch links
- Upgrading Switch-to-Server links
- Upgrading a Switched Fast Ethernet Backbone
- Upgrading a shared FDDI Backbone

This illustrated with the help of Fig. 5.6.11.

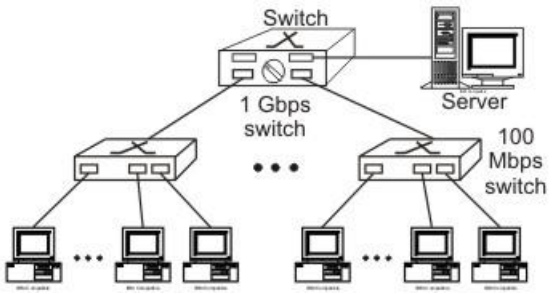


Figure 5.6.11 Migration to Gigabit Ethernet Backbone network

2.12. Internetworking Devices

Introduction

HILI subcommittee (IEEE802.1) of the IEEE identified the following possible internetworking scenarios.

- A single LAN
- Two LANs connected together (LAN-LAN)
- A LAN connected to a WAN (LAN-WAN)
- Two LANs connected through a WAN (LAN-WAN-LAN)

Various internetworking devices such as hubs, bridges, switches, routers and gateways are required to link them together. These internetworking devices are introduced in this lesson.

Repeaters

A single Ethernet segment can have a maximum length of 500 meters with a maximum of 100 stations (in a cheapernet segment it is 185m). To extend the length of the network, a *repeater* may be used. Functionally, a repeater can be considered as two transceivers joined together and connected to two different segments of coaxial cable. The repeater passes the digital signal bit-by-bit in both directions between the two segments. As the signal passes through a repeater, it is amplified and regenerated at the other end. The repeater does not isolate one segment from the other, if there is a collision on one segment, it is regenerated on the other segment. Therefore, the two segments form a single LAN and it is transparent to rest of the system. Ethernet allows five segments to be used in cascade to have a maximum network span of 2.5 km. With reference of the ISO model, a repeater is considered as a *level-1 relay* .It simply repeats, retimes and amplifies the bits it receives. The repeater is merely used to extend the span of a single LAN. Important features of a repeater are as follows:

- A repeater connects different segments of a LAN
- A repeater forwards every frame it receives
- A repeater is a regenerator, not an amplifier
- It can be used to create a single extended LAN

Hubs

Hub is a generic term, but commonly refers to a multiport repeater. It can be used to create multiple levels of hierarchy of stations. The stations connect to the hub with RJ-45 connector having maximum segment length is 100 meters. This type of interconnected set of stations is easy to maintain and diagnose. Figure 6.1.3 shows how several hubs can be connected in a hierarchical manner to realize a single LAN of bigger size with a large number of nodes.

Bridges

The device that can be used to interconnect two separate LANs is known as a bridge. It is commonly used to connect two similar or dissimilar LANs. The bridge operates in layer 2, that is data-link layer and that is why it is called level-2 relay with reference to the OSI model. It links similar or dissimilar LANs, designed to store and forward frames, it is protocol independent and transparent to the end stations. The flow of information through a bridge. Use of bridges offer a number of advantages, such as higher reliability, performance, security, convenience and larger geographic coverage. But, it is desirable that the quality of service (QOS) offered by a bridge should match that of a single LAN. The parameters that define the QOS include availability, frame mishaps, transit delay, frame lifetime, undetected bit errors, frame size and priority. Key features of a bridge are mentioned below:

• A bridge operates both in physical and data-link layer
• A bridge uses a table for filtering/routing
• A bridge does not change the physical (MAC) addresses in a frame
• Types of bridges:
    o Transparent Bridges
    o Source routing bridges

A bridge must contain addressing and routing capability. Two routing algorithms have been proposed for a bridged LAN environment. The first, produced as an extension of IEEE 802.1 and applicable to all IEEE 802 LANs, is known as transparent bridge. And the other, developed for the IEEE 802.5 token rings, is based on source routing approach. It applies to many types of LAN including token ring, token bus and CSMA/CD bus.

Transparent Bridges

The transparent bridge uses two processes known as bridge forwarding and bridge learning. If the destination address is present in the forwarding database already created, the packet is forwarded to the port number to which the destination host is attached. If it is not present, forwarding is done on all parts (flooding). This process is known as bridge forwarding. Moreover, as each frame arrives, its source address indicates where a particular host is situated, so that the bridge learns which way to forward frames to that address. This process is known as bridge learning. Key features of a transparent bridge are:

• The stations are unaware of the presence of a transparent bridge
• It performs two functions:
o Forwarding
o Learning to create th
Bridge Forwarding -Bridge forwarding operation is explain functions of the bridge forwarding
• Discard the frame if source and destination addresses are same

• Forward the frame if the source and destination

Loop Problem

Forwarding and learning processes work without any problem as long as there is no redundant bridge in the system. On the other hand, redundancy is desirable from the viewpoint of reliability, so that the function of a failed bridge is taken over by a redundant bridge. The existence of redundant bridges creates the so-called loop problem. Assuming that after initialization tables in both the bridges are empty let us consider the following steps:

Step 1. Station-A sends a frame to Station-B. Both the bridges forward the frame to LAN Y and update the table with the source address of A.

Step 2. Now there are two copies of the frame on LAN-Y. The copy sent by Bridge-a is received by Bridge-b and vice versa. As both the bridges have no information about Station B, both will forward the frames to LAN-X.

Step 3. Again both the bridges will forward the frames to LAN-Y because of the lack of information of the Station B in their database and again Step-2 will be repeated, and so on. So, the frame will continue to loop around the two LANs indefinitely.

Spanning Tree

As redundancy creates loop problem in the system, it is very undesirable. To prevent loop problem and proper working of the forwarding and learning processes, there must be only one path between any pair of bridges and LANs between any two segments in the entire bridged LAN. The IEEE specification requires that the bridges use a special topology. Such a topology is known as spanning tree (a graph where there is no loop) topology.
Source Routing Bridges

The second approach, known as source routing, where the routing operation is performed by the source host and the frame specifies which route the frame is to follow. A host can discover a route by sending a discovery frame, which spreads through the entire network using all possible paths to the destination. Each frame gradually gathers addresses as it goes. The destination responds to each frame and the source host chooses an appropriate route from these responses. For example, a route with minimum hop-count can be chosen. Whereas transparent bridges do not modify a frame, a source routing bridge adds a routing information field to the frame. Source routing approach provides a shortest path at the cost of the proliferation of discovery frames, which can put a serious extra burden on the network. Figure 6.1.11 shows the frame format of a source routing bridge.

Switches

A switch is essentially a fast bridge having additional sophistication that allows faster processing of frames. Some of important functionalities are:

• Ports are provided with buffer

• Switch maintains a directory: #address - port#
• Each frame is forwarded after examining the #address and forwarded to the proper port
• Three possible forwarding approaches: Cut-through, Collision-free and Fully-buffered as briefly explained below.

Cut-through: A switch forwards a frame immediately after receiving the destination address. As a consequence, the switch forwards the frame without collision and error detection.

Collision-free: In this case, the switch forwards the frame after receiving 64 bytes, which allows detection of collision. However, error detection is not possible because switch is yet to receive the entire frame.

Fully buffered: In this case, the switch forwards the frame only after receiving the entire frame. So, the switch can detect both collision and error free frames are forwarded.

Comparison between a switch and a hub

Although a hub and a switch apparently look similar, they have significant differences. both can be used to realize physical star topology, the hubs works like a logical bus, because the same signal is repeated on all the ports. On the other hand, a switch functions like a logical star with the possibility of the communication of separate signals between any pair of port lines. As a consequence, all the ports of a hub belong to the same collision domain, and in case of a switch each port operates on separate collision domain. Moreover, in case of a hub, the bandwidth is shared by all the stations connected to all the ports. On the other hand, in case of a switch, each port has dedicated bandwidth. Therefore, switches can be used to increase the bandwidth of a hub-based network by replacing the hubs by switches.

Routers

A router is considered as a layer-3 relay that operates in the network layer, that is it acts on network layer frames. It can be used to link two dissimilar LANs. A router isolates LANs in to subnets to manage and control network traffic. However, unlike bridges it is not transparent to end stations. A router has four basic components: Input ports, output ports, the routing processor and the switching fabric. The functions of the four components are briefly mentioned below.

• Input port performs physical and data-link layer functions of the router. As shown in
• Output ports, as shown in Fig. 6.1.14(b), perform the same functions as the input ports, but in the reverse order.
• The routing processor performs the function of the network layer. The process involves table lookup.
• The switching fabric, shown in Fig. 6.1.15, moves the packet from the input queue to the output queue by using specialized mechanisms. The switching fabric is realized with the help of multistage interconnection networks.
• Communication of a frame through a router is shown in Fig. 6.1.16.

2.13. Gateways

A gateway works above the network layer, such as application layer As a consequence, it is known as a Layer-7 relay. The application level gateways can look into the content application layer packets such as email before forwarding it to the other side. This property has made it suitable for use in Firewalls discussed in the next module.

A Simple Internet

A simple internet comprising several LANs and WANs linked with the help of routers

2.14. Frame Relay

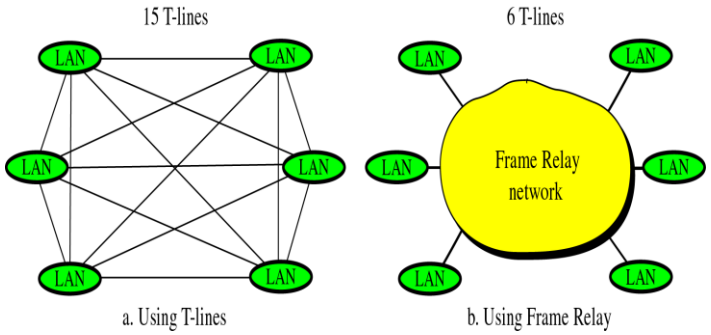


Figure 18-14

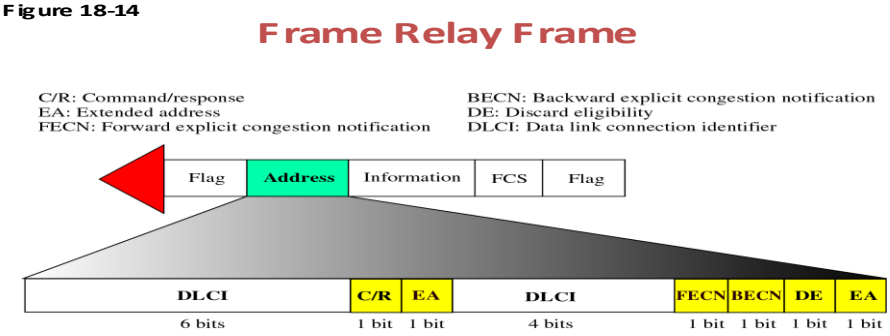
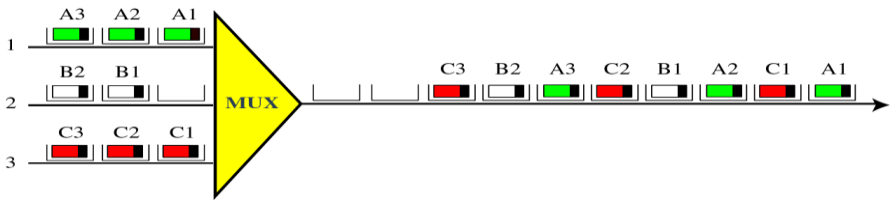


Figure 19-3

ATM Multiplexing



Introduction

Frame Relay is a high-performance WAN protocol that operates at the physical and data link layers of the OSI reference model. Frame Relay originally was designed for use across Integrated Services Digital Network (ISDN) interfaces. Today, it is used over a variety of other network interfaces as well. Frame Relay is a simplified form of Packet Switching, similar in principle to X.25, in which synchronous frames of data are routed to different destinations depending on header information. The biggest difference between Frame Relay and X.25 is that X.25 guarantees data integrity and network managed flow control at the cost of some network delays. Frame Relay switches packets end to end much faster, but there is no guarantee of data integrity at all.

As line speeds have increased from speeds below 64kbps to T1/E1 and beyond, the delays inherent in the store-and-forward mechanisms of X.25 become intolerable. At the same time, improvements in digital transmission techniques have reduced line errors to the extent that node-to-node error correction throughout the network is no longer necessary. The vast majority of Frame Relay traffic consists of TCP/IP or other protocols that provide their own flow control and error correction mechanisms. Much of this traffic is fed into the Internet, another packet switched network without any built-in error control.

Because Frame Relay does not 'care' whether the frame it is switching is error-free or not, a Frame Relay node can start switching traffic out onto a new line as soon as it has read the first two bytes of addressing information at the beginning of the frame. Thus a frame of data can travel end-to-end, passing through several switches, and still arrive at its destination with only a few bytes' delay. These delays are small enough that network latency under Frame Relay is not noticeably different from direct leased line connections. As a result, the performance of a Frame Relay network is virtually identical to that of a leased line, but because most of the network is shared, costs are lower.

Frame Relay is an example of a packet-switched technology. Packet-switched networks enable end stations to dynamically share the network medium and the available bandwidth. The following two techniques are used in packet-switching technology:

• Variable-length packets
• Statistical multiplexing

Variable-length packets are used for more efficient and flexible data transfers. These packets are switched between the various segments in the network until the destination is reached.

Statistical multiplexing techniques control network access in a packet-switched network. The advantage of this technique is that it accommodates more flexibility and more efficient use of bandwidth. Most of today's popular LANs, such as Ethernet and Token Ring, are packet-switched networks.

Frame Relay Devices

Devices attached to a Frame Relay WAN fall into the following two general categories:

• Data terminal equipment (DTE)
• Data circuit-terminating equipment (DCE)

DTEs generally are considered to be terminating equipment for a specific network and typically are located on the premises of a customer. In fact, they may be owned by the customer. Examples of DTE devices are terminals, personal computers, routers, and bridges.

DCEs are carrier-owned internetworking devices. The purpose of DCE equipment is to provide clocking and switching services in a network, which are the devices that actually transmit data through the WAN. In most cases, these are packet switches.

The connection between a DTE device and a DCE device consists of both a physical layer component and a link layer component. The physical component defines the mechanical, electrical, functional, and procedural specifications for the connection between the devices. One of the most commonly used physical layer interface specifications is the recommended standard (RS)-232 specification. The link layer component defines the protocol that establishes the connection between the DTE device, such as a router, and the DCE device, such as a switch.

Virtual Circuits

Frame Relay is a virtual circuit network, so it doesn't use physical addresses to define the DTEs connected to the network. Frame Relay provides connection-oriented data link layer communication. This means that a defined communication exists between each pair of devices and that these connections are associated with a connection identifier. However, virtual circuit identifiers in Frame relay operate at the data link layer, in contrast with X.25, where they operate at the network layer. This service is implemented by using a Frame Relay virtual circuit, which is a logical connection created between two

data terminal equipment (DTE) devices across a Frame Relay packet-switched network (PSN).

Virtual circuits provide a bidirectional communication path from one DTE device to another and are uniquely identified by a data-link connection identifier (DLCI). A number of virtual circuits can be multiplexed into a single physical circuit for transmission across the network. This capability often can reduce the equipment and network complexity required to connect multiple DTE devices.

A virtual circuit can pass through any number of intermediate DCE devices (switches) located within the Frame Relay PSN. Before going into the details of DLCI let us first have a look at the two types of Frame Relay Circuits, namely: switched virtual circuits (SVCs) and permanent virtual circuits (PVCs).

Switched virtual circuits (SVCs) are temporary connections used in situations requiring only sporadic data transfer between DTE devices across the Frame Relay network. A communication session across an SVC consists of the following four operational states:

• Call setup—The virtual circuit between two Frame Relay DTE devices is established.
• Data transfer—Data is transmitted between the DTE devices over the virtual circuit.
• Idle—The connection between DTE devices is still active, but no data is transferred. If an SVC remains in an idle state for a defined period of time, the call can be terminated.
• Call termination—The virtual circuit between DTE devices is terminated.

After the virtual circuit is terminated, the DTE devices must establish a new SVC if there is additional data to be exchanged. It is expected that SVCs will be established, maintained, and terminated using the same signaling protocols used in ISDN.

Permanent Virtual Circuits

Permanent virtual circuits (PVCs) are permanently established connections that are used for frequent and consistent data transfers between DTE devices across the Frame Relay network. Communication across PVC does not require the call setup and termination states that are used with SVCs. PVCs always operate in one of the following two operational states:

 Data transfer: Data is transmitted between the DTE devices over the virtual circuit.
Idle: The connection between DTE devices is active, but no data is transferred.

Unlike SVCs, PVCs will not be terminated under any circumstances when in an idle state. DTE devices can begin transferring data whenever they are ready because the circuit is permanently established.

Data-Link Connection Identifier (DLCI)

Frame Relay virtual circuits are identified by data-link connection identifiers (DLCIs). DLCI values typically are assigned by the Frame Relay service provider (for example, the

telephone company). Frame Relay DLCIs have local significance, which means that their values are unique in the LAN, but not necessarily in the Frame Relay WAN. The local DTEs use this DLCI to send frames to the remote DTE.

DLCIs are not only used to define the virtual circuit between a DTE and a DCE, but also to define the virtual circuit between two DCEs (switches) inside the network. A switch assigns a DLCI to each virtual connection in an interface. This means that two different connections belonging to two different interfaces may have the same DLCIs (as shown in the above figure). In other words, DLCIs are unique for a particular interface.

A connection between DTE A and DTE D has been shown in this figure, DLCI assigned inside the Frame Relay network is also shown in the network. DCEs inside the network use incoming interface – DLCI combination to decide the outgoing interface – DLCI combination to switch out the frame, from that DCE.

Each switch in a Frame relay network has a table to route frames The table matches the incoming interface- DLCI combination with an outgoing interface-DLCI combination.

Frame Relay Layers

Frame Relay has only 2 layers, namely Physical layer and Data Link layer. And as compared to other layer of packet switching network such as X.25, frame relay has only 1.5 layers whereas X.25 has 2 layers. Frame Relay eliminates all network layer functions and a portion of conventional data-link layer functions.

Physical Layer

No specific protocol is defined for physical layer in frame relay. Frame relay supports any one of the protocols recognized by ANSI, and thus the choice of physical layer protocol is up to the implementer.

Data Link Layer

At Data-link Layer Frame employs a simpler version of HDLC. Simpler version is used because HDLC provides extensive error and flow control fields that are not needed in frame relay.

To understand much of the functionality of Frame Relay, it is helpful to understand the structure of the Frame Relay frame. Figure 4.5.4 depicts the basic format of the Frame Relay frame. Flags indicate the beginning and end of the frame. Three primary

components make up the Frame Relay frame: the header and address area, the user-data portion, and the frame check sequence (FCS). The address area, which is 2 bytes in length, is comprised of 10 bits representing the actual circuit identifier and 6 bits of fields related to congestion management. This identifier commonly is referred to as the data-link connection identifier (DLCI).

• Flags—Delimits the beginning and end of the frame. The value of this field is always the same and is represented either as the hexadecimal number 7E or as the binary number 01111110.
• Address—Contains the following information:

DLCI—The 10-bit DLCI is the essence of the Frame Relay header. This value represents the virtual connection between the DTE device and the switch. Each virtual connection that is multiplexed onto the physical channel will be represented by a unique DLCI. The DLCI values have local significance only, which means that they are unique only to the physical channel on which they reside. Therefore, devices at opposite ends of a connection can use different DLCI values to refer to the same virtual connection. The first 6-bits of the first byte make up part 1 of the DLCI, and second part of DLCI uses the first 4-bits of second byte.

Extended Address (EA)—The EA is used to indicate whether the byte in which the EA value is 1 is the last addressing field. If the value is 1, then the current byte is determined to be the last DLCI octet. Although current Frame Relay implementations all use a two-octet DLCI, this capability does allow longer DLCIs to be used in the future. The eighth bit of each byte of the Address field is used to indicate the EA.

C/R—The C/R is the bit that follows the most significant DLCI byte in the Address field. The C/R bit is not currently defined.

Congestion Control—This consists of the 3 bits that control the Frame Relay congestion-notification mechanisms. These are the FECN, BECN, and DE bits, which are the last 3 bits in the Address field.

Forward-explicit congestion notification (FECN) is a single-bit field that can be set to a value of 1 by a switch to indicate to an end DTE device, such as a router, that congestion was experienced in the direction of the frame transmission from source to destination as shown in Fig. 4.5.5. The primary benefit of the use of the FECN and BECN fields is the capability of higher-layer protocols to react intelligently to these congestion indicators. Today, DECnet and OSI are the only higher-layer protocols that implement these capabilities.

Backward-explicit congestion notification (BECN) is a single-bit field that, when set to a value of 1 by a switch, indicates that congestion was experienced in the network in the direction opposite of the frame transmission from source to destination.

Discard eligibility (DE) is set by the DTE device, such as a router, to indicate that the marked frame is of lesser importance relative to other frames being transmitted. Frames

that are marked as "discard eligible" should be discarded before other frames in a congested network. This allows for a basic prioritization mechanism in Frame Relay networks.

Backward-explicit congestion notification

•Data—Contains encapsulated upper-layer data. Each frame in this variable-length field includes a user data or payload field that will vary in length up to 16,000 octets. This field serves to transport the higher-layer protocol packet (PDU) through a Frame Relay network.

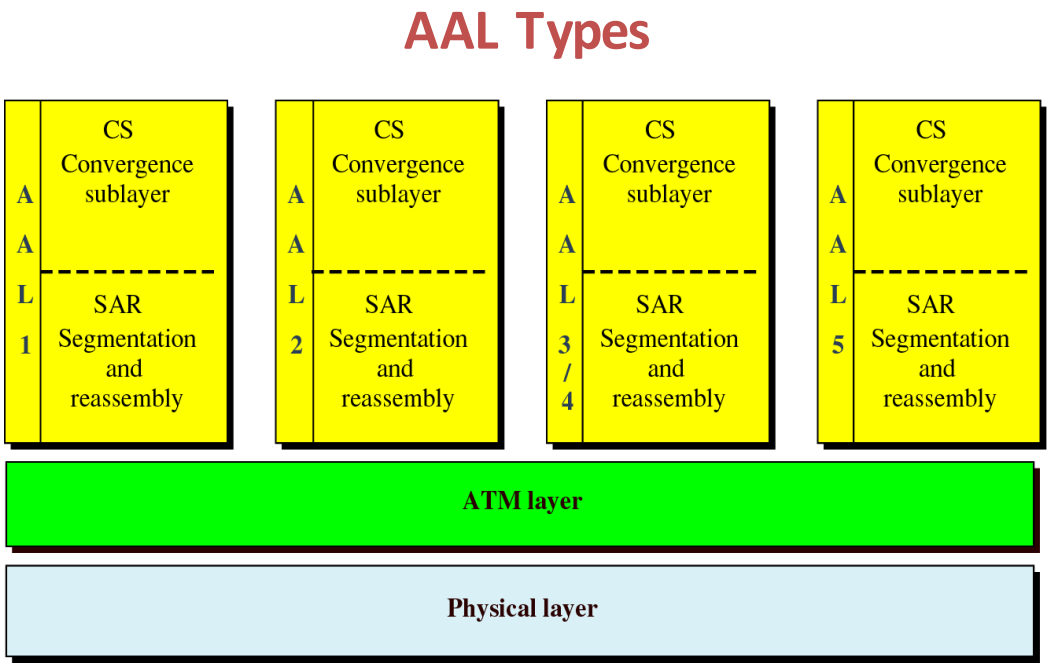
•Frame Check Sequence—Ensures the integrity of transmitted data. This value is computed by the source device and verified by the receiver to ensure integrity of transmission.

Summary

• Frame relay operates only in data link and physical layer.
• Frame Relay allows bursty traffic.
• It allows frame size of 9000 bytes, which can accommodate all local area network frames.
• Frame relay is less expensive then other traditional WANs.
• Frame relay provides both Permanent and switched connections.
• Frame relay allow variable-length frames, this may create varying delays for different users. Due to variable delay it is not suitable for real-time communication

2.15. Asynchronous Transfer Mode Switching (ATM)

Figure 19-22



Introduction

Asynchronous Transfer Mode (ATM) is an International Telecommunication Union-Telecommunications Standards Section (ITU-T) standard for cell relay wherein information for multiple service types, such as voice, video, or data, is conveyed in small, fixed-size cells. ATM networks are connection-oriented. Asynchronous transfer mode (ATM) is a technology that has its history in the development of broadband ISDN in the 1970s and 1980s. Technically, it can be viewed as an evolution of packet switching. Like packet switching protocols for data (e.g., X.25, frame relay, Transmission Control Protocol and Internet protocol (TCP IP]), ATM integrates the multiplexing and switching functions, is well suited for bursty traffic (in contrast to circuit switching), and allows communications between devices that operate at different speeds. Unlike packet switching, ATM is designed for high-performance multimedia networking. ATM technology has been implemented in a very broad range of networking devices. The most basic service building block is the ATM virtual circuit, which is an end-to-end connection that has defined end points and routes but does not have bandwidth dedicated to it. Bandwidth is allocated on demand by the network as users have traffic to transmit. ATM also defines various classes of service to meet a broad range of application needs. This lesson provides an overview of ATM protocols, services, and operation.

Benefits of ATM

The high-level benefits delivered through ATM services deployed on ATM technology using international ATM standards can be summarized as follows:

• **Dynamic bandwidth for bursty traffic** meeting application needs and delivering high utilization of networking resources; most applications are or can be viewed as inherently bursty, for example voice is bursty, as both parties are neither speaking at once nor all the time; video is bursty, as the amount of motion and required resolution varies over time.
• **Smaller header** with respect to the data to make the efficient use of bandwidth.
• **Can handle Mixed network traffic very efficiently:** Variety of packet sizes makes traffic unpredictable. All network equipments should incorporate elaborate software systems to manage the various sizes of packets. ATM handles these problems efficiently with the fixed size cell.
• **Cell network:** All data is loaded into identical cells that can be transmitted with complete predictability and uniformity.

• Class-of-service support for multimedia traffic allowing applications with varying throughput and latency requirements to be met on a single network.
• Scalability in speed and network size supporting link speeds of T1/E1 to OC–12 (622 Mbps).
• Common LAN/WAN architecture allowing ATM to be used consistently from one desktop to another; traditionally, LAN and WAN technologies have been very different, with implications for performance and interoperability. But ATM technology can be used either as a LAN technology or a WAN technology.
• International standards compliance in central-office and customer-premises environments allowing for multivendor operation.

ATM Devices and the Network Environment

ATM is a cell-switching and multiplexing technology that combines the benefits of circuit switching (guaranteed capacity and constant transmission delay) with those of packet switching (flexibility and efficiency for intermittent traffic). It provides scalable bandwidth from a few megabits per second (Mbps) to many gigabits per second (Gbps). Because of its asynchronous nature, ATM is more efficient than synchronous technologies, such as time-division multiplexing (TDM).

With TDM, each user is assigned to a time slot, and no other station can send in that time slot as shown in Fig. 4.6.1. If a station has much data to send, it can send only when its time slot comes up, even if all other time slots are empty. However, if a station has nothing to transmit when its time slot comes up, the time slot is sent empty and is wasted.

Because ATM is asynchronous, time slots are available on demand with information identifying the source of the transmission contained in the header of each ATM cell. Figure 4.6.2 shows how cells from 3 inputs have been multiplexed. At the first clock tick input 2 has no data to send, so multiplexer fills the slot with the cell from third input. When all cells from input channel are multiplexed then output slot are empty.

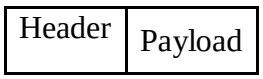
ATM Devices

An ATM network is made up of an ATM switch and ATM endpoints. An ATM switch is responsible for cell transit through an ATM network. The job of an ATM switch is well defined. It accepts the incoming cell from an ATM endpoint or another ATM switch. It then reads and updates the cell header information and quickly switches the cell to an output interface towards its destination. An ATM endpoint (or end system) contains an ATM network interface adapter. Examples of ATM endpoints are workstations, routers, digital service units (DSUs), LAN switches, and video coder-decoders (Codec's).

ATM Network Interfaces

An ATM network consists of a set of ATM switches interconnected by point-to-point ATM links or interfaces. ATM switches support two primary types of interfaces: UNI and NNI as shown in Fig. 4.6.3. The UNI (User-Network Interface) connects ATM end systems (such as hosts and routers) to an ATM switch. The NNI (Network-Network Interface) connects two ATM switches. Depending on whether the switch is owned and located at the customer's premises or is publicly owned and operated by the telephone company, UNI and NNI can be further subdivided into public and private UNIs and NNIs. A private UNI connects an ATM endpoint and a private ATM switch. Its public counterpart connects an ATM endpoint or private switch to a public switch. A private NNI connects two ATM switches within the same private organization. A public one connects two ATM switches within the same public organization.

ATM transfers information in fixed-size units called cells. Each cell consists of 53 octets, or bytes as shown in Fig. 4.6.4. The first 5 bytes contain cell-header information, and the remaining 48 contain the payload (user information). Small, fixed-length cells are well suited to transfer voice and video traffic because such traffic is intolerant to delays that result from having to wait for a large data packet to download, among other things.



An ATM cell header can be one of two formats: UNI or NNI. The UNI header is used for communication between ATM endpoints and ATM switches in private ATM networks. The NNI header is used for communication between ATM switches. Figure 4.6.5 depicts the ATM UNI cell header format, and the ATM NNI cell header format. Unlike the UNI, the NNI header does not include the Generic Flow Control (GFC) field. Additionally, the NNI header has a Virtual Path Identifier (VPI) field that occupies the first 12 bits, allowing for larger trunks between public ATM switches.

Figure 4.6.4 ATM cell Format

GFC	VPI
VPI	VCI
VPI	PT
VCI	CLP
PT	HEC
CLP	Payload
HEC	(48 bytes)
Payload	
(48 bytes)	

ATM Cell Header Fields

The following descriptions summarize the ATM cell header fields shown in Fig. 4.6.5.

• Generic Flow Control (GFC)—Provides local functions, such as identifying multiple stations that share a single ATM interface. This field is typically not used and is set to its default value of 0 (binary 0000).
• Virtual Path Identifier (VPI)—In conjunction with the VCI, identifies the next destination of a cell as it passes through a series of ATM switches on the way to its destination.
• Virtual Channel Identifier (VCI)—In conjunction with the VPI, identifies the next destination of a cell as it passes through a series of ATM switches on the way to its destination.
• Payload Type (PT)—Indicates in the first bit whether the cell contains user data or control data. If the cell contains user data, the bit is set to 0. If it contains control data, it is set to 1. The second bit indicates congestion (0 = no congestion, 1 = congestion), and the third bit indicates whether the cell is the last in a series of cells that represent a single AAL5 frame (1 = last cell for the frame).
• Cell Loss Priority (CLP)—Indicates whether the cell should be discarded if it encounters extreme congestion as it moves through the network. If the CLP bit equals 1, the cell should be discarded in preference to cells with the CLP bit equal to 0.
• Header Error Control (HEC)—Calculates checksum only on the first 4 bytes of the header. HEC can correct a single bit error in these bytes, thereby preserving the cell rather than discarding it.

ATM Virtual Connections

ATM standard defines two types of ATM connections: virtual path connections (VPCs), which contain virtual channel connections (VCCs) as shown in Fig. 4.6.6. A virtual channel connection (or virtual circuit) is the basic unit, which carries a single stream of cells, in order, from user to user. A collection of virtual circuits can be bundled together into a virtual path connection. A virtual path connection can be created from end-to-end across an ATM network. In this case, the ATM network does not route cells belonging to a particular virtual circuit. All cells belonging to a particular virtual path are routed the same way through the ATM network, thus resulting in faster recovery in case of major failures. In this

case, all the switches within the ATM network are only VP switches, i.e. they switch the cells only on the basis of VPIs. Only the switches, which are connected to the subscribers are VP/VC switches, i.e. they use both VPIs and VCIs to switch the cell. This configuration is usually followed so that the intermediate switches can do switching much faster.

Virtual channel connections of ATM

An ATM network also uses virtual paths internally for the purpose of bundling virtual circuits together between switches. Two ATM switches may have many different virtual channel connections between them, belonging to different users. These can be bundled by two ATM switches into a virtual path connection. This can serve the purpose of a virtual trunk between the two switches. This virtual trunk can then be handled as a single entity by perhaps, multiple intermediate virtual paths cross connects between the two virtual circuit switches.

ATM Switching Operations

The basic operation of an ATM switch is straightforward: The cell is received across a link with a known VPI/VCI value. The switch looks up the connection value in a local translation table to determine the outgoing port (or ports) of the connection and the new VPI/VCI value of the connection on that link. The switch then retransmits the cell on that outgoing link with the appropriate connection identifier.

Incoming		Outgoing		
VPI	VCI	VPI	VCI	Interface
10	122	11	41	1
121	213	10	158	1
12	11	211	111	2
11	151	321	210	2

A VP/VC ATM switch table

Because all VCIs and VPIs have only local significance across a particular link, these values are remapped, as necessary, at each switch. Figure 4.6.7 and Fig. 4.6.8 shows a VP-VC switch and an only VP switch, respectively. Usually the intermediate switches are only VPI switches while switches connected to the users are VPI/VCI switches.

Incoming	Outgoing	
VPI	VPI	Interface

22	65	1
121	99	2
312	201	1
11	21	2

VP ATM switch table

To make the switching more efficient, ATM uses two types of switches namely, VP switch and VP-VC switch. A VP switch route cells only on the basis of VPI, here VPIs change but VCIs remain same during switching. On the other hand, VP-VC switch uses the complete identifier, i.e. both VPI and VCI to route the cell. We can think of a VP-VC switch as a combination of Only VP and Only VC switch.

ATM Reference Model

The ATM architecture uses a logical model to describe the functionality that it supports. ATM functionality corresponds to the physical layer and part of the data link layer of the OSI reference model.

The ATM reference model, as shown in Fig. 4.6.9, consists of the following planes, which span all layers:

• Control—This plane is responsible for generating and managing signaling requests.
• User—This plane is responsible for managing the transfer of data.
• Management—This plane contains two components:
Layer management manages layer-specific functions, such as the detection of failures and protocol problems.
Plane management manages and coordinates functions related to the complete system.

The ATM reference model consists of the following ATM layers:

• Physical layer—Analogous to the physical layer of the OSI reference model, the ATM physical layer manages the medium-dependent transmission.

• ATM layer—Combined with the ATM adaptation layer, the ATM layer is roughly analogous to the data link layer of the OSI reference model. The ATM layer is responsible for the simultaneous sharing of virtual circuits over a physical link (cell multiplexing) and passing cells through the ATM network (cell relay). To do this, it uses the VPI and VCI information in the header of each ATM cell.
• ATM adaptation layer (AAL)—Combined with the ATM layer, the AAL is roughly analogous to the data link layer of the OSI model. The AAL is responsible for isolating higher-layer protocols from the details of the ATM processes. The adaptation layer prepares user data for conversion into cells and segments the data into 48-byte cell payloads.

Finally, the higher layers residing above the AAL accept user data, arrange it into packets, and hand it to the AAL.

The ATM Physical Layer

The main functions of the ATM physical layer are as follows:

• Cells are converted into a bit stream,
• The transmission and receipt of bits on the physical medium are controlled,
• ATM cell boundaries are tracked,
• Cells are packaged into the appropriate types of frames for the physical medium.

The ATM physical layer is divided into two parts: the physical medium-dependent (PMD) sub layer and the transmission convergence (TC) sub layer.

The PMD sub layer provides two key functions.

• It synchronizes transmission and reception by sending and receiving a continuous flow of bits with associated timing information.
• It specifies the physical media for the physical medium used, including connector types and cable.

The TC sub layer has four functions:

• Cell delineation, it maintains ATM cell boundaries, allowing devices to locate cells within a stream of bits.
• Generates and checks the header error control code to ensure valid data.
• Cell-rate decoupling, maintains synchronization and inserts or suppresses idle (unassigned) ATM cells to adapt the rate of valid ATM cells to the payload capacity of the transmission system.
• Transmission frame adaptation packages ATM cells into frames acceptable to the particular physical layer implementation.

ATM Layer

The ATM layer provides routing, traffic management, switching and multiplexing services. It processes outgoing traffic by accepting 48-byte segment from the AAL sub-layers and transforming them into 53-byte cell by addition of a 5-byte header. Adaptation Layers

ATM adaptation layers allow existing packet networks to connect to ATM facilities. AAL Protocol accepts transmission from upper layer services (e.g.: packet data) and map them into fixed-sized ATM cells. These transmissions can be of any type, variable or fixed data rate. At the receiver, this process is reversed and segments are reassembled into their original formats and passed to the receiving services. Instead of one protocol for all types of data, the ATM standard divides the AAL layer into categories, each supporting the requirements of different types of applications. There are four types of data streams that are identified: Constant-bit rate, variable bit-rate, connection oriented packet data transfer, connectionless packet data transfer. In addition to dividing AAL by category (AAL1, AAL2 and so on), ITU-T also divides it on the basis of functionality. Each AAL layer is actually divided into two layers: the convergence sub-layer and Segmentation and reassembly (SAR) sub-layer. Table below gives a brief description of these data streams and various ATM adaptation layers which are used for each of them.

Table Mapping of various data types and ATM adaptation layers

Service Class	Quality of Service Parameter	ATM Adaptation layers

Constant Bit rate (CBR)	This class is used for emulating circuit switching. The cell rate is constant with time. CBR applications are quite sensitive to cell-delay variation. Examples of applications that can use CBR are telephone traffic (i.e., nx64 kbps), videoconferencing, and television.	AAL1: AAL1, a connection-oriented service, is suitable for handling constant bit rate sources (CBR), such as voice and videoconferencing. AAL1 requires timing synchronization between the source and the destination. For this reason, AAL1 depends on a medium, such as SONET, that supports clocking. The AAL1 process prepares a cell for transmission in three steps. First, synchronous samples (for example, 1 byte of data at a sampling rate of 200 microseconds) are inserted into the Payload field. Second, Sequence Number (SN) and Sequence Number Protection (SNP) fields are added to provide information that the receiving AAL1 uses to verify that it has received cells in the correct order. Third, the remainder of the Payload field is filled with enough single bytes to equal 48 bytes.

Variable Bit Rate - non-real time (VBR–NRT)	This class allows users to send traffic at a rate that varies with time depending on the availability of user information. Statistical multiplexing is provided to make optimum use of network resources. Multimedia e-mail is an example of VBR–NRT.	AAL 2: The AAL2 process uses 44 bytes of the cell payload for user data and reserves 4 bytes of the payload to support the AAL2 processes. VBR traffic is characterized as either real-time (VBR-RT) or as non-real-time (VBR-NRT). AAL2 supports both types of VBR traffic.
Variable bit rate–real time (VBR–RT)	This class is similar to VBR–NRT but is designed for applications that are sensitive to cell-delay variation. Examples for real-time VBR are voice with speech activity detection (SAD) and interactive compressed video.	

Connection oriented packet transfer or available bit rate (ABR)	This class of ATM services provides rate-based flow control and is aimed at data traffic such as file transfer and e-mail. Although the standard does not require the cell transfer delay and cell-loss ratio to be guaranteed or minimized, it is desirable for switches to minimize delay and loss as much as possible. Depending upon the state of congestion in the network, the source is required to control its rate. The users are allowed to declare a minimum cell rate, which is guaranteed to the connection by the network.	AAL3/4: AAL3/4 supports both connection-oriented and connectionless data. AAL3/4 prepares a cell for transmission in four steps. First, the convergence sub layer (CS) creates a protocol data unit (PDU) by prepending a beginning/end tag header to the frame and appending a length field as a trailer. Second, the segmentation and reassembly (SAR) sub layer fragments the PDU and prepends a header to it. Then the SAR sub layer appends a CRC-10 trailer to each PDU fragment for error control. Finally, the completed SAR PDU becomes the Payload field of an ATM cell to which the ATM layer prepends the standard ATM header. AAL 5: AAL5 is the primary AAL for data and supports both connection-oriented and connectionless data. It is used to transfer most non-SMDS data, such as classical IP over ATM and LAN Emulation (LANE). AAL5 also is known as the simple and efficient adaptation layer (SEAL)
Connectionless data transfer or unspecified bit rate (UBR)	This class is the catch-all, other class and is widely used today for TCP/IP.	

ATM Applications

ATM is used in both LANs and WANs; let's have a look at few of the possible applications.

ATM WANs: ATM is basically a WAN technology that delivers cell over long distances. Here ATM is mainly used to connect LANs or other WANs together. A router between ATM network and the other network serves as an end point. This router has two stacks of protocols: one belonging to ATM and other belonging to other protocol.

ATM LANs: High data rate (155 and 622 Mbps) of ATM technology attracted designers to think of implementing ATM technology in LANs too. At the surface level, to implement an ATM LAN ATM switch will replace the traditional Ethernet switch, in a switched LAN. But few things have to be kept in mind and software modules would be needed to map the following differences between the two technologies:

• Connectionless versus connection-oriented: ATM is a virtual connection oriented technology, while traditional Ethernet uses connectionless protocols.
• Physical address versus virtual circuit identifier: In the Traditional LAN packets are routed based on the source and destination addresses, while in ATM cells are routed based on the virtual circuit identifiers (VPI-VCI pair).

Multimedia virtual private networks and managed services: Service providers are building on their ATM networks to offer a broad range of services. Examples include managed ATM, LAN, voice and video services (these being provided on a per-application basis, typically including customer-located equipment and offered on an end-to-end basis), and full-service virtual private-networking capabilities (these including integrated multimedia access and network management).

Frame-relay backbones: Frame-relay service providers are deploying ATM backbones to meet the rapid growth of their frame-relay services to use as a networking infrastructure for a range of data services and to enable frame relay to ATM service internetworking services.

Internet backbones: Internet service providers are likewise deploying ATM backbones to meet the rapid growth of their frame-relay services, to use as a networking infrastructure for a range of data services, and to enable Internet class-of-service offerings and virtual private intranet services.

Residential broadband networks: ATM is the networking infrastructure of choice for carriers establishing residential broadband services, driven by the need for highly scalable solutions.

Carrier infrastructures for the telephone and private-line networks: Some carriers have identified opportunities to make more-effective use of their SONET/SDH fiber infrastructures by building an ATM infrastructure to carry their telephony and private-line traffic.

Short question Answers:

Q-1. What is the advantage of token passing protocol over CSMA/CD protocol?

Ans. Advantage of token passing protocol over CSMA/CD protocol:
The CSMA/CD is not a deterministic protocol. A packet may be delivered after many (up to 15) collisions leading to long variable delay. An unfortunate packet may not get delivered at all. This feature makes CSMA/CD protocol unsuitable for real-time applications. On the other hand, token passing protocol is a deterministic approach, which allows a packet to be delivered within a known time frame. It also allows priority to be assigned to packets. These are the two key advantages of token passing protocol over CSMA/CD protocol.

Q-2. What are the drawbacks of token ring topology?

Ans. Token ring protocol cannot work if a link or a station fails. So, it is vulnerable to link and station failure.

Q-3. How the reliability of token ring topology can be improved?

Ans. Reliability of the ring network can be improved by implementing the ring topology using a wiring concentrator. This allows not only to detect fault, but also to isolate the faulty link/station with the help of a bypass relay.

Q-4. What role the active token monitor performs?

Ans. Token ring is maintained with the help of active token monitor. Any one of the stations has the capability to act as active token monitor, but at a particular instant only

Q-5. Explain the basic difference between IEEE 802.3 and switched Ethernet, as far as implementation is concerned.

Ans: In Ethernet (IEEE 802.3) the topology, though physically is start but logically is BUS. i.e. the collision domain of all the nodes in a LAN is common. In this situation only one frame can send the

frame, if more than one station sends the frame, there is a collision. In Switched Ethernet, this collision domain is separated. Hub is replaced by a switch, a device that can recognize the destination address and can route the frame to the port to which the destination station is connected, the rest of the media is not involved in the transmission process. The switch can receive another frame from another station at the same time and can route this frame to its own final destination.

Q-6. Explain the two techniques for implementing Ethernet switches.

Ans: There are two techniques used in the implementation of Ethernet switches: *store-and-forward* and *cut-through*. In the first case, the entire frame is captured at the incoming port, stored in the switch's memory, and after an address lookup to determine the LAN destination port, forwarded to the appropriate port. The lookup table is automatically built up. On the other hand, a cut-through switch begins to transmit the frame to the destination port as soon as it decodes the destination address from the frame header.

Store-and-forward approach provides a greater level of error detection because damaged frames are not forwarded to the destination port. But, it introduces longer delay of about 1.2 msec for forwarding a frame and suffers from the chance of loosing data due to reliance on buffer memory. The cut-through switches, on the other hand, has reduced latency but has higher switch cost.

Q-7. What are the different categories of Fast Ethernet?

Ans: IEEE has designed two categories of Fast Ethernet: 100Base-X and 100Base-T4. 100Base-X uses two cables between hub and the station while 100Base-T4 uses four. 100-Base-X itself is divided into two: 100Base-TX and 100base-FX.

* *100 BASE-T4*: This option is designed to avoid overwriting. It is used for half-duplex communication using four wire-pairs of the existing category 3 UTP cable, which is already available for telephone services in homes/offices. Two of four pairs are bi-directional; other two are unidirectional. This means that there are 3 pairs to be used for carrying data, in each direction (2 bi-directional and 1 uni-directional). Because 100Mbps data cannot be handled by voice-grade UTP, this specification splits the 100 Mbps flow into three 33.66Mbps flow.

**100 BASE TX*: This option uses two category 5 UTP or two shielded (STP) cable to connect a station to hub. One pair is used to carry frames from the hub to the station and other to carry frames from station to hub. Encoding is 4B/5B to handle 100 Mbps; signaling is NRZ-I. The distance between station and hub should be less than 100 meters.

**100 BASE FX*: This option uses two Fiber optic cables, one carry frames from station to hub and other from hub to station. The encoding is 4B/5B and signaling in NRZ-I. the distance between station and hub should be less than 2000 meters.

Q-8. What are the Objectives of The Gigabit Ethernet Alliance?

Ans: The objectives of the alliance are:

• supporting extension of existing Ethernet and Fast Ethernet technology in response to demand for higher network bandwidth.
• developing technical proposals for the inclusion in the standard

• establishment of inter-operability test procedures and processes

Q-9. Explain GMII (Gigabit Media Independent Interface) in brief.

Ans: The GMII is the interface between the MAC layer and the Physical layer. It allows any physical layer to be used with the MAC layer. It is an extension of the MII (Media Independent Interface) used in Fast Ethernet. It uses the same management interface as MII. It supports 10, 100 and 1000 Mbps data rates. It provides separate 8-bit wide receive and transmit data paths, so it can support both full-duplex as well as half-duplex operation.

----------x----------x----------

UNIT III

NETWORK LAYER

Logical addressing: IPv4, IPv6 addresses Internet Protocol: Internetworking – IPv4, IPv6 - Address mapping – ARP, RARP, BOOTP, DHCP, ICMP, IGMP, Delivery - Forwarding - Routing – Unicast, Multicast routing protocols

---

Specific Instructional Objectives

At the end of this lesson, the students will be able to:

- Explain the relationship between TCP/IP and OSI model
- Explain different classes of IP addresses
- Explain the concept of subnetting and subnet masking
- Explain the ARP/RARP protocol
- Explain fragmentation and reassembly
- Explain the ICMP protocols
- State the key features of IPv6

3.1 Introduction

In the previous lesson we have discussed various devices required for internetworking. In addition to these devices, several protocols are required to provide necessary functionality for internetworking. The software that provide these protocols is known as Transmission Control Protocol/Internet Protocol (TCP/IP). TCP/IP acts as a glue to link different types of LAN and WAN to provide Internet, a single integrated network for seamless communication. The IP provides unreliable, connectionless best-effort

datagram delivery service, whereas TCP provides reliable, efficient and cost-effective end-to-end delivery of data. The relationship between TCP/IP and the OSI model is shown in Fig. 6.2.1. This lesson introduces the IP protocol and various issues related to it.

3.2 Addressing

To send a packet from a source node to a destination node correctly through a network, the packet must contain enough information about the destination address. It is also common to include the source address, so that retransmission can be done, if necessary. The addressing scheme used for this purpose has considerable effect on routing.

There are two possible approaches used for addressing; *flat* and *hierarchical*. In *flat addressing* every possible node is assigned a unique number. When a new node is added to the network, it must be given an address within the allowed address range. Addressing used in Ethernet is an example of flat addressing, where addresses (48-bits long) are allocated centrally, blocks of addresses are apportioned to manufactures, so that no two devices in the world will have the same address. Flat addressing has the advantage that if a node is moved from one location to another, it can retain its unique address.

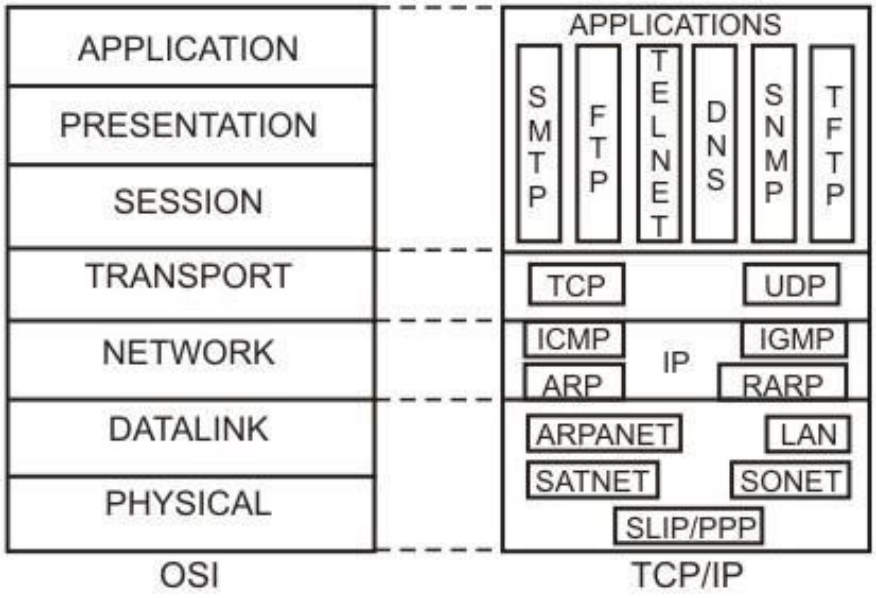


Figure 6.2.1 Relationship between the TCP/IP and the OSI model

In *hierarchical addressing*, each address consists of a number of fields; as each field is inspected, the packet is taken nearer to the destination. This is very similar to the addressing used in postal system. A significant advantage of hierarchical addressing is that it is possible to relate a hierarchical address structure to the topology of the network, so that routing is simplified. This scheme has the disadvantage that if a host moves from one location to another, a new address needs to be allocated to it, in the same manner that an address change is required as we change house.

3.3 IP Addressing

Every host and router on the internet is provided with a unique standard form of network address, which encodes its network number and host number. The combination is unique;no two nodes have the same IP addresses. The IP addresses are 32-bit long having the formats shown in Fig 6.2.2. The three main address formats are assigned with network addresses (net id) and host address (host id) fields of different sizes. The class A format allows up to 126 networks with 16 million hosts each. Class B allows up to 16,382 networks with up to 64 K hosts each. Class C allows 2 million networks with up to 254 hosts each. The Class D is used for multicasting in which a datagram is directed to multiple hosts. Addresses beginning with 11110 are reserved for future use. Network addresses are usually written in dotted decimal notation, such as 126.12.15.220, where each byte is written in decimal number corresponding to the binary value. Figure 6.2.3 illustrates how the dotted decimal representation is obtained for a particular IP address in binary form. Range of IP addresses for different classes is given in Fig. 6.2.4. Some IP addresses, which are used in special situations such as the same host, a host the same network, broadcast on the same network, broadcast on a distant network, or loopback are given in Fig. 6.2.5. This approach of representing IP addresses in terms of classes is known as *classful addressing*. In mid 90's another approach known as *classlees*

addressing has been proposed, which may supersede the existing classful addressing approach in future.

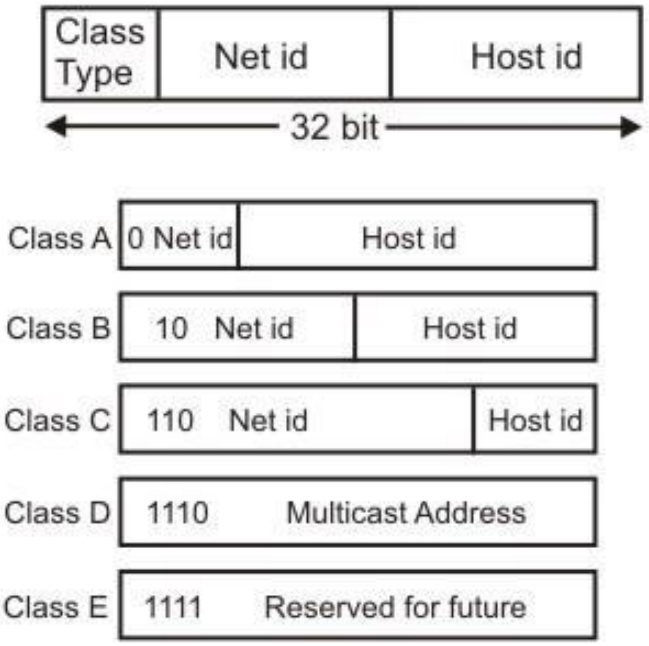


Figure 6.2.2 IP address formats

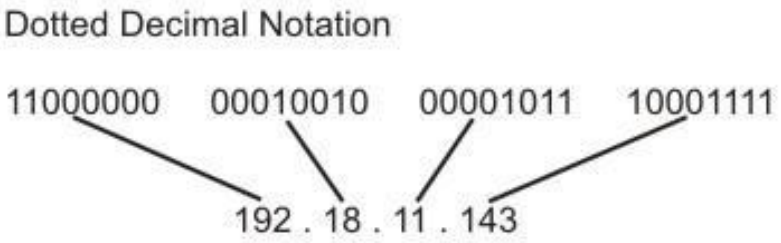


Figure 6.2.3 Dotted decimal representation

Range of Host Addresses		
Class A	1.0.0.0	to 127.255.255.255
Class B	128.0.0.0	to 191.255.255.255
Class C	192.0.0.0	to 233.255.255.255
Class D	244.0.0.0	to 239.255.255.255
Class E	240.0.0.0	to 247.255.255.255

Figure 6.2.4 Dotted decimal notation of the IP addresses

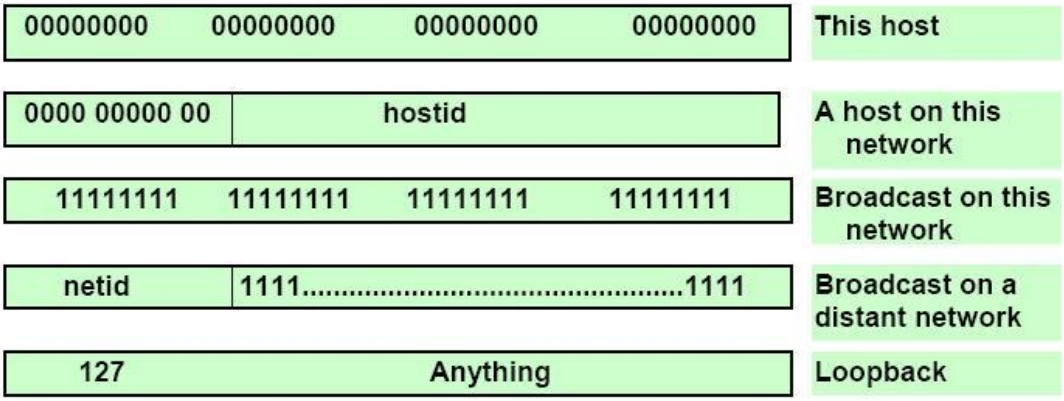


Figure 6.2.5 Special IP addresses

3.4 Subnetting

To filter packets for a particular network, a router uses a concept known as *masking*, which filters out the net id part (by ANDing with all 1's) by removing the host id part (by
ANDing with all 0's). The net id part is then compared with the network address as shown in Fig. 6.2.6. All the hosts in a network must have the same network number. This property of IP addressing causes problem as the network grows. To overcome this problem, a concept known as *subnets* is used, which splits a network into several parts for internal use, but still acts like a single network to the outside world. To facilitate routing, a concept known as *subnet mask* is used. As shown in Fig. 6.2.7, a part of hostid is used as subnet address with a corresponding subnet mask. Subnetting reduces router table space by creating a three-level hierarchy; net id, subnet id followed by hosted.

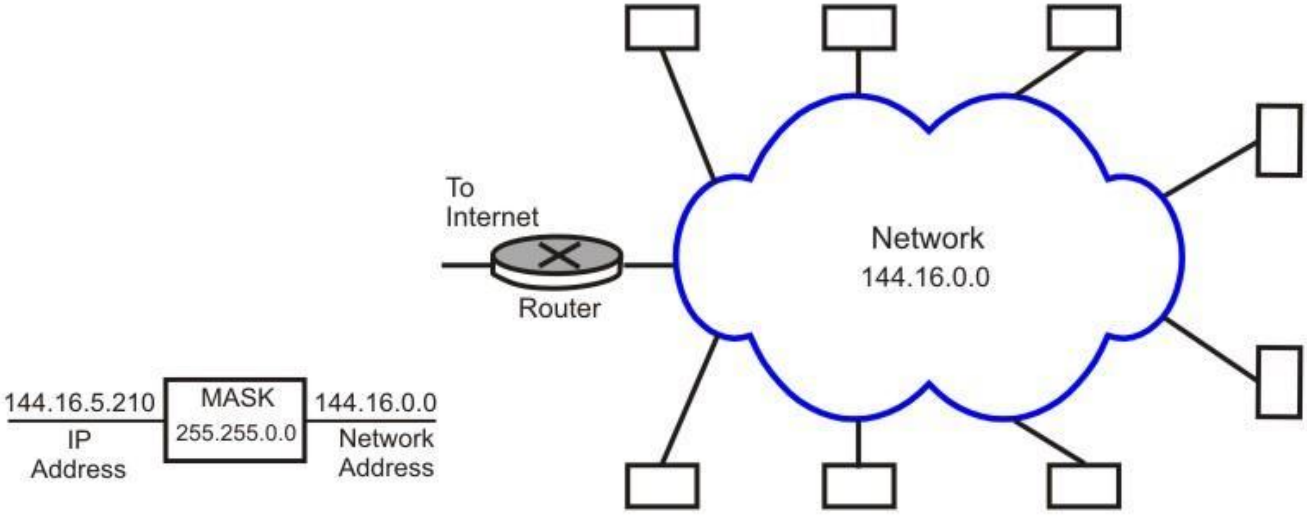


Figure 6.2.6 Masking with the help of router

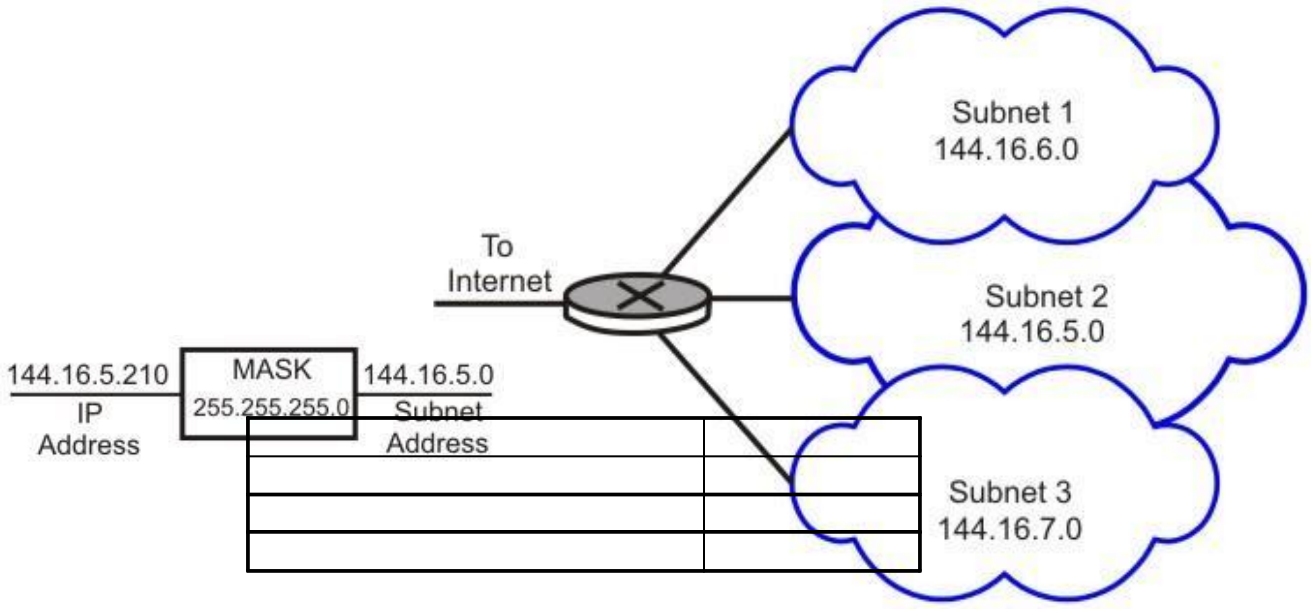


Figure 6.2.7 Subnet masking with the help of router

3.5 Network Address Translation (NAT)

With the increasing number of internet users requiring an unique IP address for each host, there is an acute shortage of IP addresses (until everybody moves to IPV6). The *Network Address Translation* (NAT) approach is a quick interim solution to this problem. NAT allows a large set of IP addresses to be used in an internal (private) network and a handful of addresses to be used for the external internet. The internet authorities has set aside three sets of addresses to be used as private addresses as shown in Table 6.2.1. It may benoted that these addresses can be reused within different internal networks simultaneously, which in effect has helped to increase the lifespan of the IPV4. However, to make use of the concept, it is necessary to have a router to perform the operation of address translation between the private network and the internet. As shown in Fig. 6.2.8, the NAT router maintains a table with a pair of entries for private and internet address. The source address of all outgoing packets passing through the NAT router gets replaced by an internet address based on table look up. Similarly, the destination address of all incoming packets passing through the NAT router gets replaced by the corresponding private address, as shown in the figure. The NAT can use a pool of internet addresses to have internet access by a limited number of stations of the private network at a time.

Table 6.2.1 Addresses for Private Network

Range of addresses	Total number
10.0.0.0 to 10.255.255.255	2^{24}
172.16.0.0 to 172.31.255.255	2^{20}
192.168.0.0 to 192.168.255.255	2^{16}

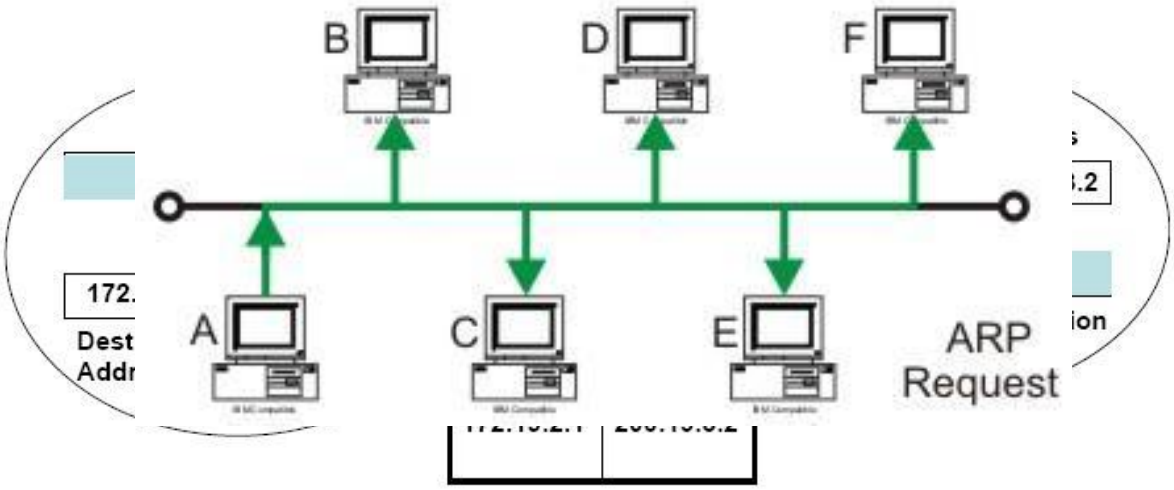


Figure 6.2.8 NAT Address translation

3.6 Address Resolution Protocol (ARP)

It may be noted that the knowledge of hosts' IP address is not sufficient for sending packets, because *data link hardware does not understand internet addresses*.For example, in an Ethernet network, the Ethernet controller card can send and receive using 48-bit Ethernet addresses. The 32-bit IP addresses are unknown to these cards. This requires a mapping of the IP addresses to the corresponding Ethernet addresses. This mapping is accomplished by using a technique known as *Address Resolution Protocol (ARP)*.

One possible approach is to have a *configuration file* somewhere in the system that maps IP addresses onto the Ethernet addresses. Although this approach is straightforward, maintaining an up-to0-date table has a high overhead on the system.Another elegant approach is to broadcast packet onto the Ethernet asking "*who owns the destination IP address*?". The destination node responds with its Ethernet address after hearing the request. This protocol of asking the question and getting the reply is called ARP (Addressing Resolution Protocol), which is widely used. ARP is a dynamic mapping approach for finding a physical address for a known IP address. It involves following two basic steps as shown in Fig. 6.2.9.

- An ARP request is broadcast to all stations in the network
- An ARP reply is an unicast to the host requesting the mapping

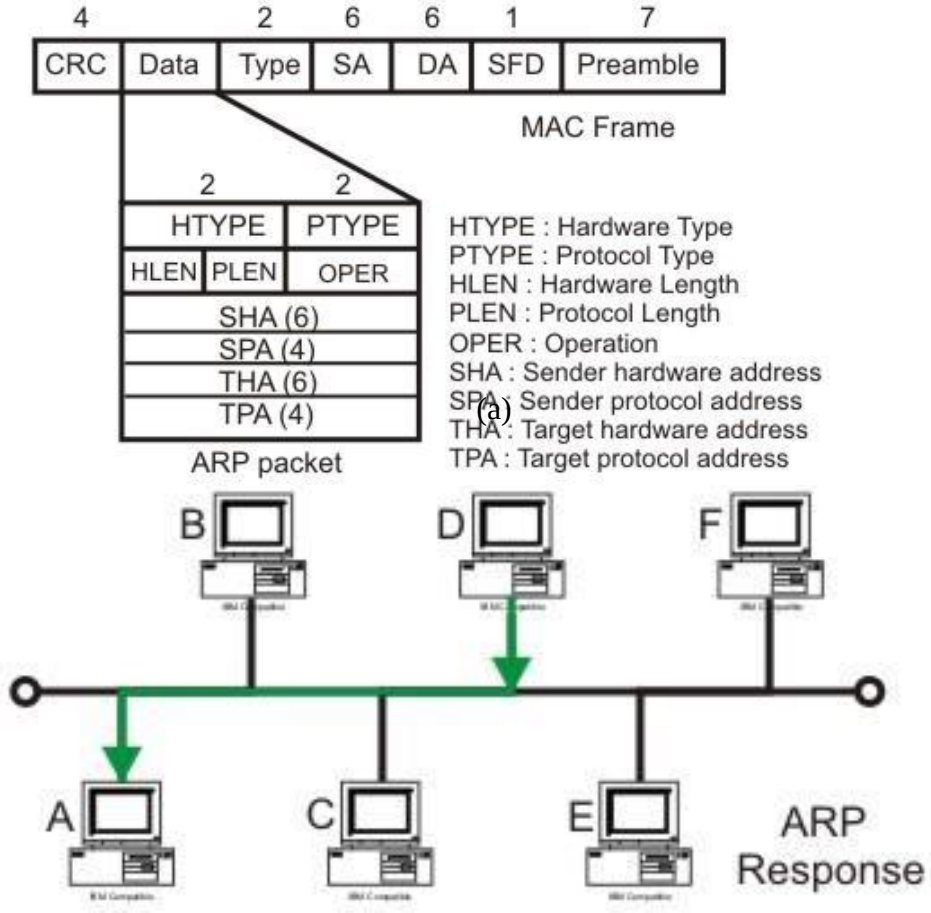


Figure 6.2.9 (a) ARP request with a broadcast to all the stations and (b) ARP response is a unicast only to the requesting host

Various optimizations are commonly used to improve the efficiency of the ARP protocol. One possible approach is to use cache memory to hold the recently acquired frame containing the physical address. As a consequence, no broadcasting is necessary in near future. Figure 6.2.10 shows how an ARP packet is encapsulated into the data field of a MAC frame.

Reverse ARP (RARP)

The TCP/IP protocols include another related protocol known as reverse ARP, which can be used by a computer such as a diskless host to find out its own IP address. It involves the following steps:

- Diskless host A broadcasts a RARP request specifying itself as the target
- RARP server responds with the reply directly to host A
- Host A preserves the IP address in its main memory for future use until it reboots

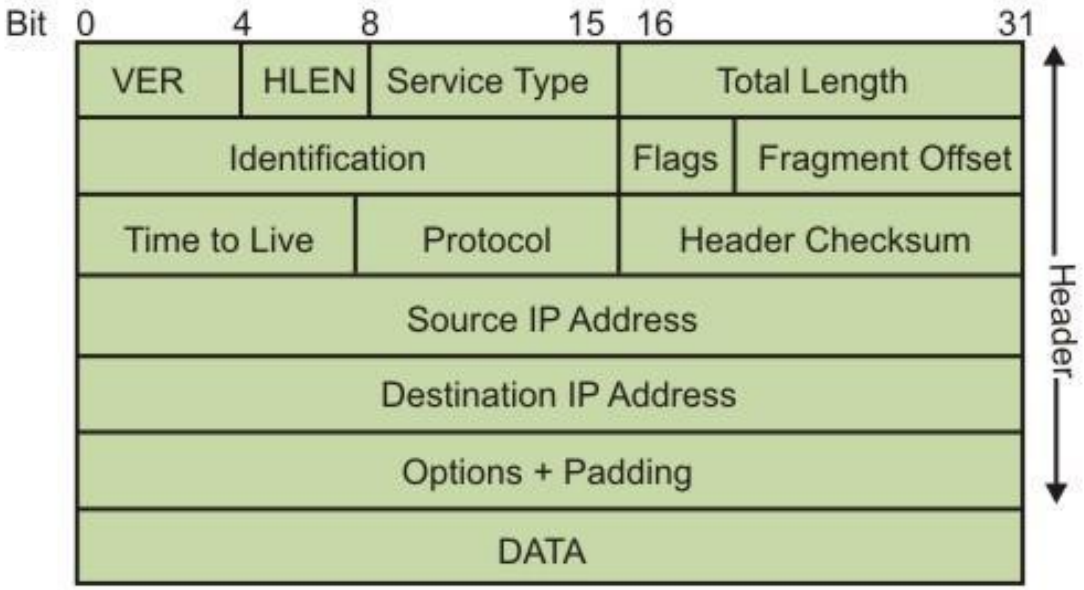


Figure 6.2.10 An ARP packet is encapsulated directly into the data field a MAC frame

3.7 IP Datagram

As we have mentioned earlier, IP is an unreliable and connectionless *best-effort* delivery service protocol. By best effort we mean that there is no error and flow control. However, IP performs error detection and discards a packet, if it is corrupted. To achieve reliability, it is necessary to combine it with a reliable protocol such as TCP. Packets in IP layer are called *datagrams*. The IP header provides information about various functions the IP performs. The IP header format is shown in Fig. 6.2.11. The 20 to 60 octets of header has a number of fields to provide:

- Source and destination IP addresses
- Non transparent fragmentation
- Error checking
- Priority
- Security
- Source routing option
- Route Recording option
- Stream identification
- Time stamping

A brief description of each of the fields are given below:

- VER (4 bits): Version of the IP protocol in use (typically 4).
- HLEN (4 bits): Length of the header, expressed as the number of 32-bit words. Minimum size is 5, and maximum 15.
- Total Length (16 bits): Length in bytes of the datagram, including headers. Maximum datagram size is (216) 65536 bytes.

- Service Type (8 bits): Allows packet to be assigned a priority. Router can use this field to route packets. Not universally used.
- Time to Live (8 bits): Prevents a packet from traveling forever in a loop. Senders sets a value, that is decremented at each hop. If it reaches zero, packet is discarded.
- Protocol: Defines the higher level protocol that uses the service of the IP layer
- Source IP address (32 bits): Internet address of the sender.
- Destination IP address (32 bits): Internet address of the destination.
- Identification, Flags, Fragment Offset: Used for handling fragmentation.
- Options (variable width): Can be used to provide more functionality to the IP datagram
- Header Checksum (16 bits):

    - Covers only the IP header.
    - Steps:
        - Header treated as a sequence of 16-bit integers
        - The integers are all added using ones complement arithmetic
        - Ones complement of the final sum is taken as the checksum
        - Datagram is discarded in case of mismatch in checksum values

3.8 Multiplexing and Demultiplexing

IP datagram can encapsulate data from several higher-level protocols such as TCP, UDP, ICMP, etc. The Protocol field in the datagram specifies the final destination protocol to which IP datagram to be delivered. When the datagram arrives at the destination, the information in this field is used to perform demultiplex the operation. The multiplexing and demultiplexing operations are shown in Fig. 6.2.12.

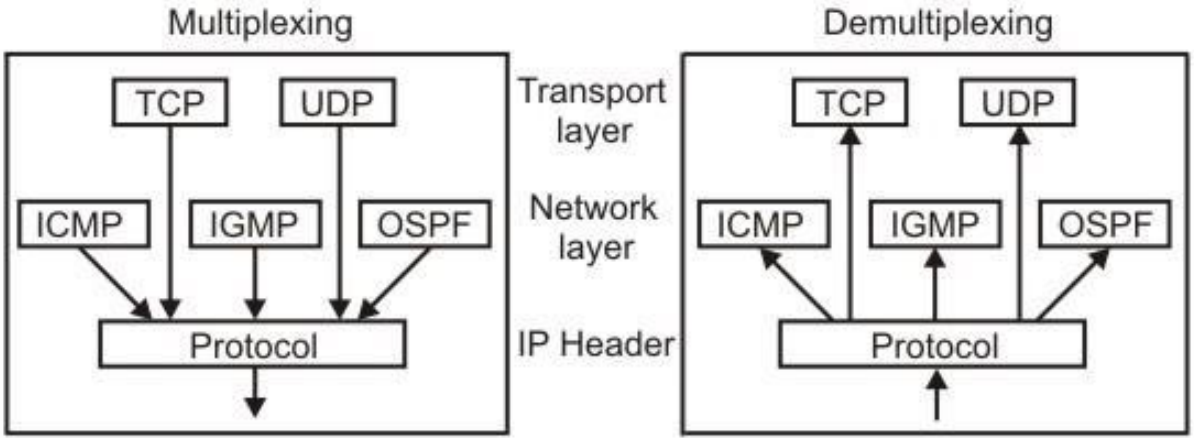


Figure 6.2.12 Multiplexing and demultiplexing in the IP layer

3.9 Fragmentation and Reassembly

Each network imposes a limit on maximum size, known as *maximum transfer unit* (MTU) of a packet because of various reasons. One approach is to prevent the problem to occur in the first place, i.e. send packets smaller than the MTU. Second approach is to deal with the problem using fragmentation. When a gateway connects two networks that have different maximum and or minimum packet sizes, it is necessary to allow the gateway to break packets up into fragments, sending each one as an internet packet. The technique is known as *fragmentation*. The following fields of an IP datagram are related to fragmentation:

- **Identification**: A16-bit field identifies a datagram originating from the source host.
- **Flags**: There are 3 bits, the first bit is reserved, the second bit is *do not fragment* bit, and the last bit *is more fragment* bit.
- **Fragmentation offset**: This 13-bit field shows the relative position of the segment with respect to the complete datagram measured in units of 8 bytes.

Figure 6.2.13 shows a fragmentation example, where a packet is fragmented into packets of 1600 bytes. So, the offset of the second fragmented packet is 1600/8 = 200 and the offset of the third fragmented packet is 400 and so on.

The reverse process, known as *reassembly*, which puts the fragments together, is a more difficult task. There are two opposing strategies for performing the re-assembly. In the first case, the fragmentation in one network is made transparent to any subsequent networks. This requires that packets to be reassembled before sending it to subsequent networks as shown in Fig. 6.2.14(a). This strategy is used in ATM. As re-assembly requires sufficient buffer space for storage of all the fragments, this approach has large storage overhead. To overcome this problem in the second strategy, re-assembly is done only at the ultimate destination. This approach does not require large buffer but additional fields are to be added to each packet for independent addressing and to indicate the fragment number as shown in Fig. 6.2.14(b).

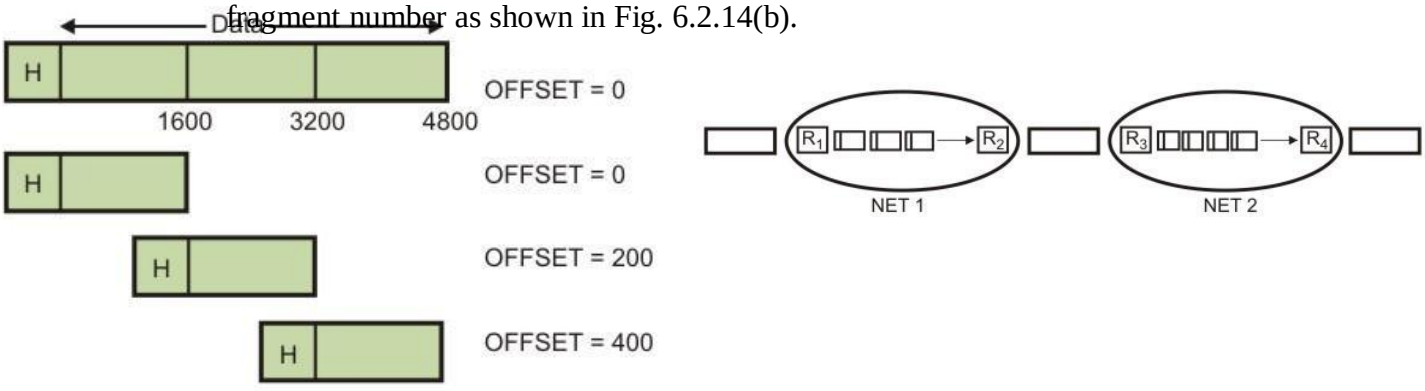


Figure 6.2.13 Fragmentation example

3.10 ICMP

To make efficient use of the network resources, IP was designed to provide unreliable and connectionless best-effort datagram delivery service. As a consequence, IP has no error-control mechanism and also lacks mechanism for host and management queries. A companion protocol known as *Internet Control Message Protocol* (ICMP), has been designed to compensate these two deficiencies. ICMP messages can be broadly divided into two broad categories: error reporting messages and query messages as follows.

- Error reporting Messages: Destination unreachable, Time exceeded, Source quench, Parameter problems, Redirect
- Query: Echo request and reply, Timestamp request and reply, Address mask request and reply

The frame formats of these query and messages are shown in Fig. 6.2.15.

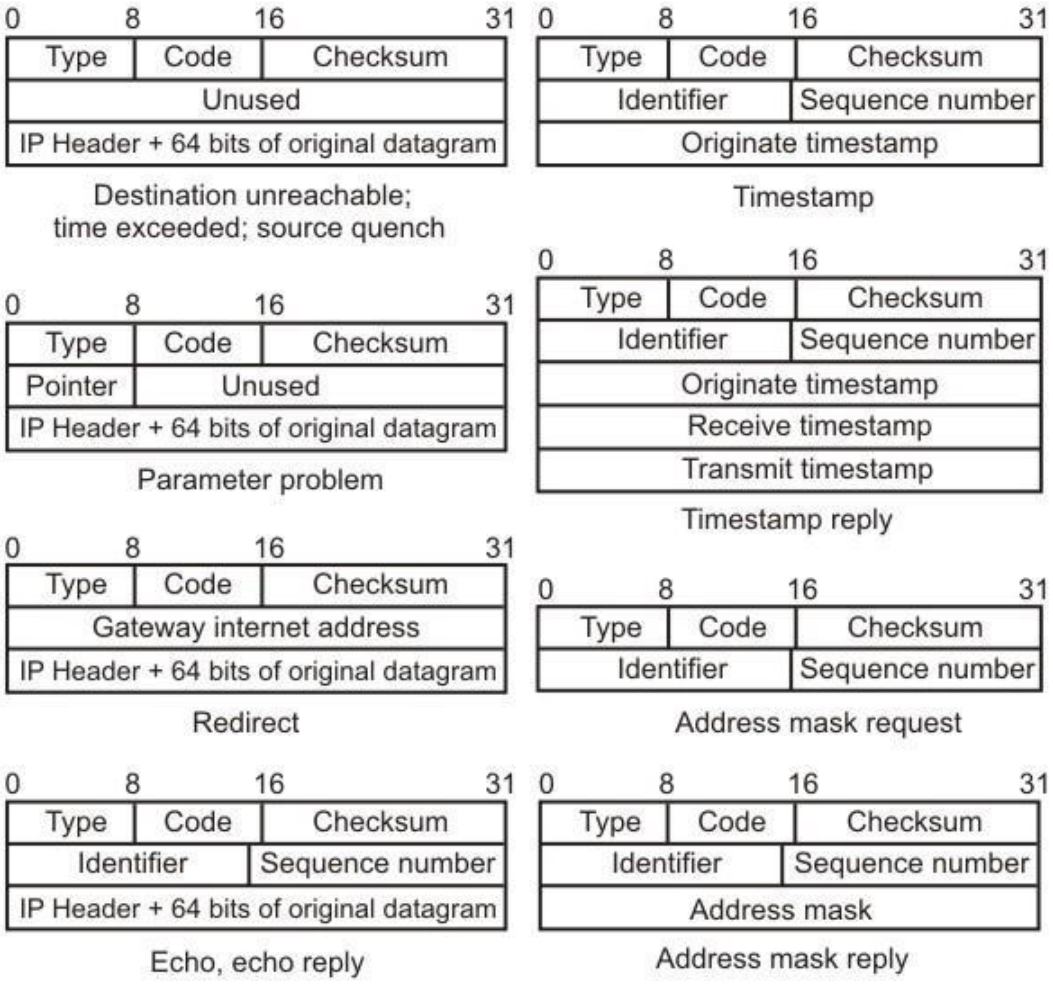


Figure 8.2. .ICMP Query and Message Formats

3.11 IPV6

The network layer that is present in use in commonly referred to as IPv4. Although IPv4 is well designed and has helped the internet to grow rapidly, it has some deficiencies, These deficiencies has made it unsuitable for the fast growing internet. To overcome these deficiencies, Internet Protocol, Version 6 protocol has been proposed and it has evolved into a standard. Important features of IPv6 are highlighted below:

- IPv6 uses 128-bit address instead of 32-bit address to provide larger address space
- Uses more flexible header format, which simplifies and speeds up the routing process
- Basic header followed by extended header
- Resource Allocation options, which was not present in IPv4
- Provision of new/future protocol options
- Support for security with the help of encryption and authentication
- Support for fragmentation at source

TWO MARKS QUESTIONS

Q1.Why do you need ARP protocol?

Ans: Two machines on a network can communicate only if they know each other's physical address. So, IP address is not enough to deliver a packet to the destination node.

It is necessary to know its physical (LAN) address. The ARP protocol allows a host to find out the physical address of a destination host on the same physical network, given only the IP address of the destination host.

Q2.What is the purpose of dotted decimal representation? Give dotted decimal representation of the IP address 11011101 10001111 11111101 00001111.

Ans: To represent the 32-bit IP address in short and easy to read form, Internet addresses are represented in decimal form with decimal points separating the bytes. This is known as dotted decimal notation. For the given IP address the dotted decimal representation is 221.143.253.15.

Q3.How is masking is related to subnetting?

Ans: Masking is a process that extracts the physical network address part from the 32-bit IP address. When subnetting is done, the masking is performed to get the subnetwork address rather than the network address.

Q4. What is the function of NAT?

Ans: The *Network Address Translation* (NAT) approach is a quick interim solution to this problem of acute shortage of IP addresses for individual hosts in IPv4. NAT allows a large set of IP addresses to be used in an internal (private) network and a handful of addresses to be used for the global internet.

Q5. What is the function of the ICMP?

Ans: The ICMP has been designed as companion protocol to compensate two important deficiencies of the IP protocol, namely error-control mechanism and the lack of mechanism for host and management queries.

UNIT IV

TRANSPORT LAYER

Process-to-Process delivery - User Datagram Protocol (UDP) – Transmission Control Protocol (TCP) – Congestion Control – Quality of services (QoS) – Techniques to improve QoS.

Specific Instructional Objectives

On completion of this lesson, the students will be able to:

- Explain the cause for congestion
- Understand the effects of congestion
- Understand various open-loop and close-loop congestion control techniques:
  - The leaky bucket algorithm
  - The token bucket algorithm
  - Admission Control
  - Choke packets
  - Weighted fair queuing
  - Load shedding
  - Resource reservation
- Distinguish between flow and congestion control

4.1 Introduction

As Internet can be considered as a *Queue of packets*, where transmitting nodes are constantly adding packets and some of them (receiving nodes) are removing packets from the queue. So, consider a situation where too many packets are present in this queue (or internet or a part of internet), such that constantly transmitting nodes are pouring packets at a higher rate than receiving nodes are removing them. This degrades the performance, and such a situation is termed as *Congestion*. Main reason of congestion is more number of packets into the network than it can handle. So, the objective of congestion control can be summarized as to maintain the number of packets in the network below the level at which performance falls off dramatically. The nature of a Packet switching network can be summarized in following points:

- A network of queues
- At each node, there is a queue of packets for each outgoing channel
- If packet arrival rate exceeds the packet transmission rate, the queue size grows without bound
- When the line for which packets are queuing becomes more than 80% utilized, the queue length grows alarmingly When the number of packets dumped into the network is within the carrying capacity, they all are delivered, expect a few that have too be

rejected due to transmission errors). And then the number delivered is proportional to the number of packets sent. However, as traffic increases too far, the routers are no longer able to cope, and they begin to lose packets. This tends to make matter worse. At very high traffic, performance collapse completely, and almost no packet is delivered. In the following sections, the causes of congestion, the effects of congestion and various congestion control techniques
are discussed in detail.

4.2 Causes Of Congestion

Congestion can occur due to several reasons. For example, if all of a sudden a stream of packets arrive on several input lines and need to be out on the same output line,then a long queue will be build up for that output. If there is *insufficient memory* to hold these packets, then packets will be lost (dropped). Adding more memory also may not help in certain situations. If router have an infinite amount of memory even then instead of congestion being reduced, it gets worse; because by the time packets gets at the head of the queue, to be dispatched out to the output line, they have already timed-out (repeatedly), and duplicates may also be present. All the packets will be forwarded to next router up to the destination, all the way only increasing the load to the network more and more. Finally when it arrives at the destination, the packet will be discarded, due to time out, so instead of been dropped at any intermediate router (in case memory is restricted) such a packet goes all the way up to the destination, increasing the network load throughout and then finally gets dropped there.

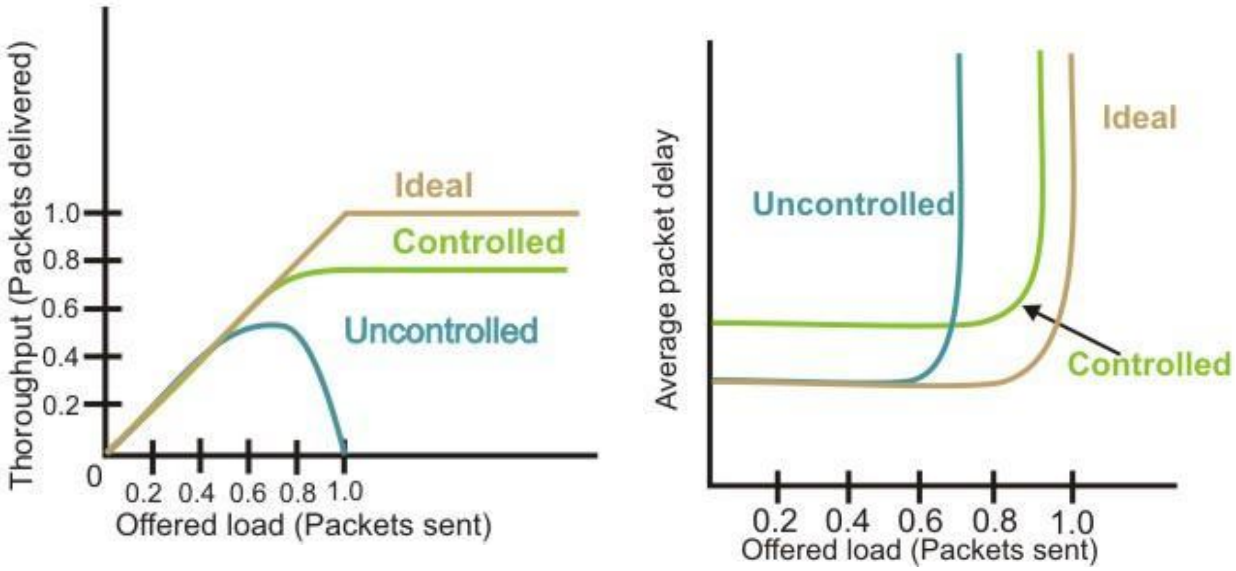
Slow processors also cause Congestion. If the router CPU is slow at performing the task required for them (Queuing buffers, updating tables, reporting any exceptions etc.), queue can build up even if there is excess of line capacity. Similarly, *Low-Bandwidth* lines can also cause congestion. Upgrading lines but not changing slow processors, or vice-versa, often helps a little; these can just shift the bottleneck to some other point. The real problem is the mismatch between different parts of the system.

Congestion tends to feed upon itself to get even worse. Routers respond to overloading by dropping packets. When these packets contain TCP segments, the segments don't reach their destination, and they are therefore left unacknowledged, which eventually leads to timeout and retransmission. So, the major cause of congestion is often the *bursty* nature of traffic. If the hostscould be made to transmit at a uniform rate, then congestion problem will be less common and all other causes will not even led to congestion because other causes just actas an enzyme which boosts up the congestion when the traffic is bursty (i.e., other causes just add on to make the problem more serious, main cause is the bursty traffic). This means that when a device sends a packet and does not receive an acknowledgment from the receiver, in most the cases it can be assumed that the packets have been dropped by intermediate devices due to congestion. By detecting the rate at which segments are sent and not acknowledged, the source or an intermediate router can infer the level of congestion on the network. In the following section we shall discuss the ill effects of congestion.

4.3 Effects of Congestion

Congestion affects two vital parameters of the network performance, namely *throughput*and *delay*. In simple terms, the throughput can be defined as the percentage utilization of the network capacity. Figure 7.5.1(a) shows how throughput is affected as offered load increases. Initially throughput increases linearly with offered load, because utilization ofthe network increases. However, as the offered load increases beyond certain limit, say 60% of the capacity of the network, the throughput drops. If the offered load increases further, a point is reached when not a single packet is delivered to any destination, which is commonly known as *deadlock* situation. There are three curves in Fig. 7.5.1(a), the ideal one corresponds to the situation when all the packets introduced are delivered to their destination up to the maximum capacity of the network. The second one corresponds to the situation when there is no congestion control. The third one is the case when some congestion control technique is used. This prevents the throughput collapse, but provides lesser throughput than the ideal condition due to overhead of the congestion control technique.

The delay also increases with offered load, as shown in Fig. 7.5.1(b). And no matter what technique is used for congestion control, the delay grows without bound as the load approaches the capacity of the system. It may be noted that initially there is longer delay when congestion control policy is applied. However, the network without any congestion control will saturate at a lower offered load.



(a) (b)

Figure 7.5.1 (a) Effect of congestion on throughput (b) Effect of congestion on delay

4.4 Congestion Control Techniques

Congestion control refers to the mechanisms and techniques used to control congestion and keep the traffic below the capacity of the network. As shown in Fig. 7.5.2, the congestion control techniques can be broadly classified two broad categories:

- **Open loop:** Protocols to prevent or avoid congestion, ensuring that the system (or network under consideration) never enters a Congested State.
- **Close loop:** Protocols that allow system to enter congested state, detect it, and remove it.

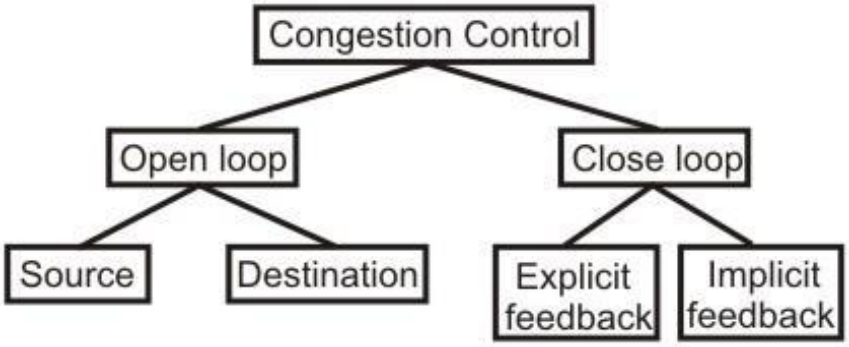


Figure 7.5.2 Congestion control categories

The first category of solutions or protocols attempt to solve the problem by a good design, at first, to make sure that it doesn't occur at all. Once system is up and running mid course corrections are not made. These solutions are somewhat static in nature, as the policies to control congestion don't change much according to the current state of the system. Such Protocols are also known as *Open Loop* solutions. These rules or policies include deciding upon when to accept traffic, when to discard it, making scheduling decisions and so on. Main point here is that they make decision without taking into consideration the current state of the network. The open loop algorithms are further divided on the basis of whether these acts on source versus that act upon destination.

The second category is based on the concept of feedback. During operation, some system parameters are measured and feed back to portions of the subnet that can take action to reduce the congestion. This approach can be divided into 3 steps:

- Monitor the system (network) to detect whether the network is congested or not and what's the actual location and devices involved.
- To pass this information to the places where actions can be taken
- Adjust the system operation to correct the problem.

These solutions are known as *Closed Loop* solutions**.** Various Metrics can be used to monitor the network for congestion. Some of them are: the average queue length, number of packets that are timed-out, average packet delay, number of packets discarded due to

lack of buffer space, etc. A general feedback step would be, say a router, which detects the congestion send special packets to the source (responsible for the congestion) announcing the problem. These extra packets increase the load at that moment of time, but are necessary to bring down the congestion at a later time. Other approaches are also used at times to curtail down the congestion. For example, hosts or routers send out probe packets at regular intervals to explicitly ask about the congestion and source itself regulate its transmission rate, if congestion is detected in the network. This kind of approach is a *pro-active* one, as source tries to get knowledge about congestion in the network and act accordingly.

Yet another approach may be where instead of sending information back to the source an intermediate router which detects the congestion send the information about the congestion to rest of the network, piggy backed to the outgoing packets. This approach will in no way put an extra load on the network (by not sending any kind of special packet for feedback). Once the congestion has been detected and this information has been passed to a place where the action needed to be done, then there are two basic approaches that can overcome the problem. These are: either to increase the resources or to decrease the load. For example, separate dial-up lines or alternate links can be used to increase the bandwidth between two points, where congestion occurs. Another example could be to decrease the rate at which a particular sender in transmitting packets out into the network.

The closed loop algorithms can also be divided into two categories, namely *explicit feedback* and *implicit feedback* algorithms. In the explicit approach, special packets are sent back to the sources to curtail down the congestion. While in implicit approach, the source itself acts pro-actively and tries to deduce the existence of congestion by making local observations.

In the following sections we shall discuss about some of the popular algorithms from the above categories.

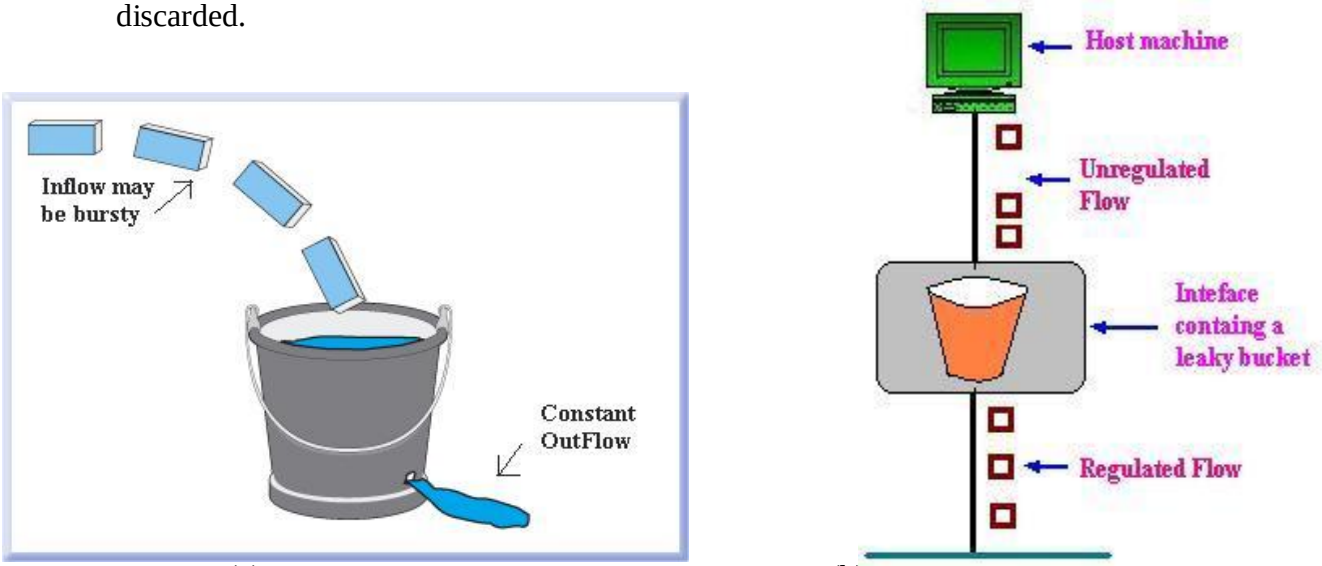
4.5 Leaky Bucket Algorithm

Consider a Bucket with a small hole at the bottom, whatever may be the rate of water pouring into the bucket, the rate at which water comes out from that small hole is constant. This scenario is depicted in figure 7.5.3(a). Once the bucket is full, any additional water entering it spills over the sides and is lost (i.e. it doesn't appear in the output stream through the hole underneath). The same idea of leaky bucket can be applied to packets, as shown in Fig. 7.5.3(b). Conceptually each network interface contains a *leaky bucket*. And the following steps are performed:

- When the host has to send a packet, the packet is thrown into the bucket.
- The bucket leaks at a constant rate, meaning the network interface transmits packets at a constant rate.
- Bursty traffic is converted to a uniform traffic by the leaky bucket.
- In practice the bucket is a finite queue that outputs at a finite rate.

This arrangement can be simulated in the operating system or can be built into the hardware.

Implementation of this algorithm is easy and consists of a finite queue. Whenever a packet arrives, if there is room in the queue it is queued up and if there is no room then the packet is discarded.



(a) (b)

Figure 7.5.3(a) Leaky bucket (b) Leaky bucket implementation

4.6 Token Bucket Algorithm

The leaky bucket algorithm described above, enforces a rigid pattern at the output stream, irrespective of the pattern of the input. For many applications it is better to allow the output to speed up somewhat when a larger burst arrives than to loose the data. Token Bucket algorithm provides such a solution. In this algorithm leaky bucket holds token, generated at regular intervals. Main steps of this algorithm can be described as follows:

- In regular intervals tokens are thrown into the bucket.
- The bucket has a maximum capacity.
- If there is a ready packet, a token is removed from the bucket, and the packet is send.
- If there is no token in the bucket, the packet cannot be send.

Figure 7.5.4 shows the two scenarios before and after the tokens present in the bucket have been consumed. In Fig. 7.5.4(a) the bucket holds two tokens, and three packets are waiting to be sent out of the interface, in Fig. 7.5.4(b) two packets have been sent out by consuming two tokens, and 1 packet is still left.

The token bucket algorithm is less restrictive than the leaky bucket algorithm, in a sense that it allows bursty traffic. However, the limit of burst is restricted by the number of tokens available in the bucket at a particular instant of time.

The implementation of basic token bucket algorithm is simple; a variable is used just to count the tokens. This counter is incremented every *t seconds* and is decremented whenever a packet is sent. Whenever this counter reaches zero, no further packet is sent out as shown in Fig. 7.5.5.

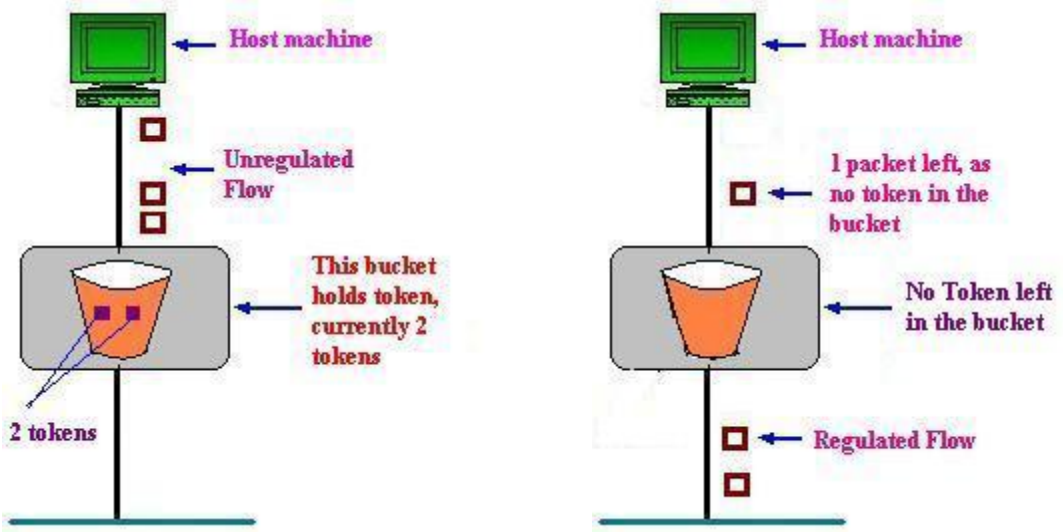


Figure 7.5.4(a) Token bucket holding two tokens, before packets are send out, (b) Token bucket after two packets are send, one packet still remains as no token is left

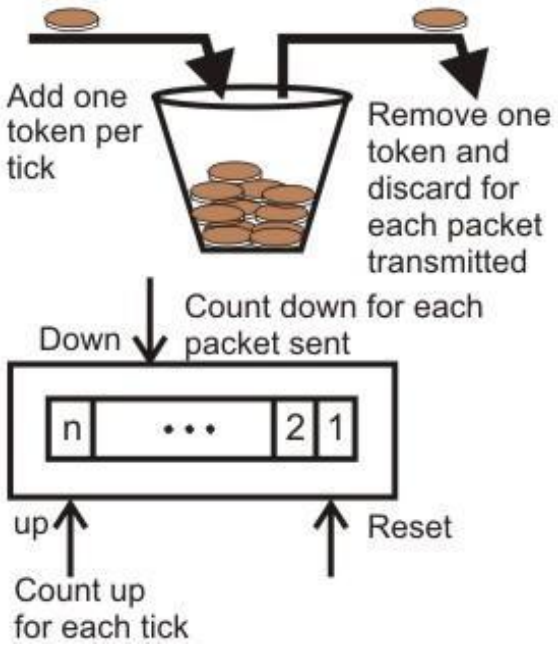


Figure 7.5.5 Implementation of the Token bucket algorithm

4.7 Congestion control in virtual Circuit

Till now we have discussed two open loop algorithms, where the policy decisions are made in the beginning, irrespective of the current state. Both leaky bucket algorithm and token bucket algorithm are open loop algorithms.

- Simpler one being: do not set-up new connections, once the congestion is signaled. This type of approach is often used in normal telephone networks. When the exchange is overloaded, then no new calls are established.
- Another approach, which can be followed is: to allow new virtual connections, but route these carefully so that none of the congested router (or none of the problem area) is a part of this route.
- Yet another approach can be: To negotiate different parameters between the host and the network, when the connection is setup. During the setup time itself, Host specifies the volume and shape of traffic, quality of service, maximum delay and other parameters, related to the traffic it would be offering to the network. Once the host specifies its requirement, the resources needed are reserved along the path, before the actual packet follows.

4.8 Choke Packet Technique

The *choke packet* technique, a closed loop control technique**,** can be applied in both virtual circuit and datagram subnets. Each router monitors its resources and the utilization at each of its output line. There is a threshold set by the administrator, and whenever any of the resource utilization crosses this threshold and action is taken to curtail down this. Actually each output line has a utilization associated with it, and whenever this utilization crosses the threshold, the output line enters a "warning" state. If so, the router sends a *choke packet* back to the source, giving it a feedback to reduce the traffic. And the original packet is tagged (a bit is manipulated in the header field) so that it will not generate other choke packets by other intermediate router, which comes in place and is forwarded in usual way. It means that the first router (along the way of a packet), which detects any kind of congestion, is the only one that sends the choke packets.

When the source host gets the choke packet, it is required to reduce down the traffic send out to that particular destination (choke packet contains the destination to which the original packet was send out). After receiving the choke packet the source reduces the traffic by a particular fixed percentage, and this percentage decreases as the subsequent choke packets are received. Figure 7.5.6 depicts the functioning of choke packets.

For Example, when source A receives a choke packet with destination B at first, it will curtail down the traffic to destination B by 50%, and if again after affixed duration of time interval it receives the choke packet again for the same destination, it will further curtail down the traffic by 25% more and so on. As stated above that a source will entertain another subsequent choke packet only after a fixed interval of time, not before that. The reason for this is that when the first choke packet arrives at that point of time other packets destined to the same destination would also be there in the network and they will generate other choke packets too, the host should ignore these choke packets which refer to the same destination for a fixed time interval.

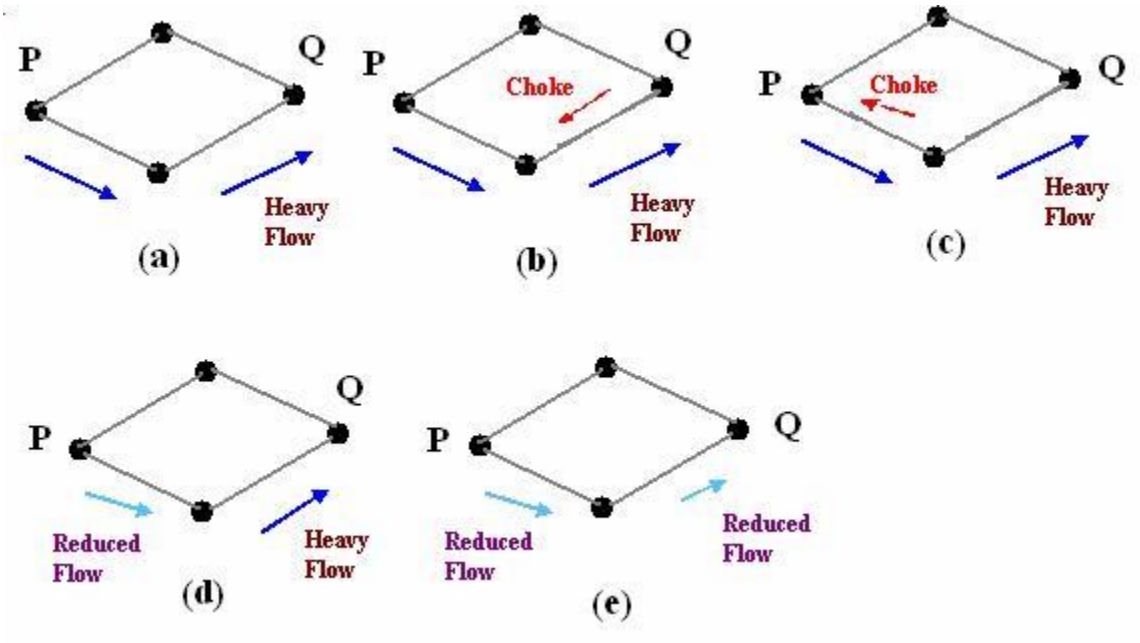


Figure 7.5.6 Depicts the functioning of choke packets, (a) Heavy traffic between nodes P and Q, (b) Node Q sends the Choke packet to P, (c) Choke packet reaches P, (d) P reduces the flow and send a reduced flow out, (e) Reduced flow reaches node Q

4.9 Hop-by Hop Choke Packets

This technique is an advancement over Choked packet method. At high speed over long distances, sending a packet all the way back to the source doesn't help much, because by the time choke packet reach the source, already a lot of packets destined to the same original destination would be out from the source. So to help this, Hop-by-Hop Choke packets are used. In this approach, the choke packet affects each and every intermediate router through which it passes by. Here, as soon as choke packet reaches a router back to its path to the source, it curtails down the traffic between those intermediate routers. In this scenario, intermediate nodes must dedicate few more buffers for the incoming traffic as the outflow through that node will be curtailed down immediately as choke packet arrives it, but the input traffic flow will only be curtailed down when choke packet reaches the node which is before it in the original path. This method is illustrated in Fig. 7.5.7.

As compared to choke packet technique, hop-by-hop choke packet algorithm is able to restrict the flow rapidly. As can been seen from Figures 7.5.6 and 7.5.7, one-step eduction is seen in controlling the traffic, this single step advantage is because in our example there is only one intermediate router. Hence, in a more complicated network, one can achieve a significant advantage by using hop-by-hop choke packet method.

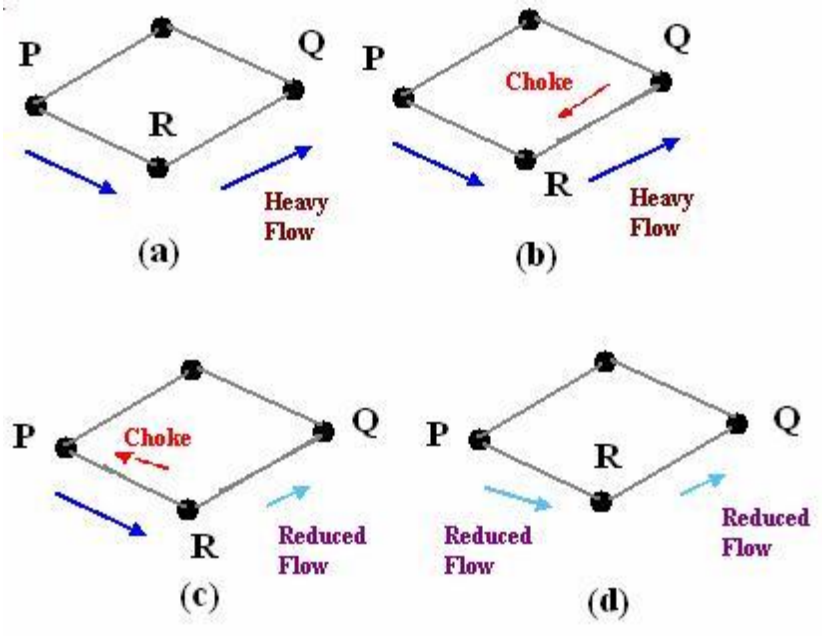


Figure 7.5.7 Depicts the functioning of Hop-by-Hop choke packets, (a) Heavy traffic between nodes P and Q, (b) Node Q sends the Choke packet to P, (c) Choke packet reaches R, and the flow between R and Q is curtail down, Choke packer reaches P, and P reduces the flow out

4.10 Load Shedding

Another simple closed loop technique is *Load Shedding***;** it is one of the simplest and more effective techniques. In this method, whenever a router finds that there is congestion in the network, it simply starts dropping out the packets. There are different methods by which a host can find out which ackets to drop. Simplest way can be just choose the packets randomly which has to be dropped. More effective ways are there but they require some kind of cooperation from the sender too. For many applications, some packets are more important than others. So, sender can mark the packets in priority
classes to indicate how important they are. If such a priority policy is implemented than intermediate nodes can drop packets from the lower priority classes and use the available bandwidth for the more important packets.

4.11 Slow Start - a Pro-active technique

This is one of the pro-active techniques, which is used to avoid congestion. In the original implementation of TCP, as soon as a connection was established between two devices, they could each go “hog wild”, sending segments as fast as they liked as long as there was room in the other devices receive window. In a busy internet, the sudden appearance of a large amount of new traffic could aggravate any existing congestion.

To alleviate this, modern TCP devices are restrained in the rate at which they initially send

segments. Each sender is at first restricted to sending only an amount of data equal to one "full-sized" segment—that is, equal to the MSS (maximum segment size) value for the connection. Each time an acknowledgment is received, the amount of data the device can send is increased by the size of another full-sized segment. Thus, the device "starts slow" in terms of how much data it can send, with the amount it sends increasing until either the full window size is reached or congestion is detected on the link. In the latter case, the congestion avoidance feature is used.

When potential congestion is detected on a TCP link, a device responds by throttling back the rate at which it sends segments. A special algorithm is used that allows the device to drop the rate at which segments are sent quickly when congestion occurs. The device then uses the *Slow Start* algorithm just above to gradually increase the transmission rate back up again to try to maximize throughput without congestion occurring again.

4.12 Flow Control Versus Congestion control

Let's have a look at the difference between *Flow Control* and *Congestion Control*, which are mixed up at times.

Flow control is a very important part of regulating the transmission of data between devices, but it is limited in a way that it only considers what is going on within each of the devices on the connection, and not what is happening in devices between them. It relates to the point-point traffic between a given sender and a receiver. Flow control always involves some kind of feedback from receiver to sender to tell sender how things are at other end of the network. Since we are dealing with how TCP works between a typical server and client at layer four, we don't worry about how data gets between them; that's the job of the Internet Protocol at layer three.

In practice, what is going on at layer three can be quite important. Considered from an abstract point of view, our server and client may be connected "directly" using TCP, but all the packets we transmit are carried across an internet and routers between different networks. These networks and routers are also carrying data from many other connections and higher-layer protocols. If the internet becomes very busy, the speed at which segments are carried between the endpoints of our connection will be reduced, and they could even be dropped. This is called *congestion control*. Congestion control has to do with making sure that subnet carry the offered traffic. It is the global issue, involving the behavior of all the hosts, router, link, store and forward mechanism between them in the entire subnet or internet.

TWO MARKS QUESTIONS

1. What are the two basic mechanisms of congestion control?

Ans : The two basic mechanisms of congestion control are:

- One is preventive, where precautions are taken so that congestion can not occur.
- Another is recovery from congestion, when congestion has already taken place

2. How congestion control is performed by leaky bucket algorithm?

Ans : In **leaky bucket algorithm,** a buffering mechanism is introduced between the host computer and the network in order to regulate the flow of traffic. Busty traffic are generated by the host computer and introduced in the network by leaky bucket mechanism in the following manner

- Packets are introduced in the network in one per tick
- In case of buffer overflow packets are discarded

3. In what way token bucket algorithm is superior to leaky bucket algorithm?

Ans : The leaky bucket algorithm controls the rate at which the packets are introduced in the network, but it is very conservative in nature. Some flexibility is introduced in token bucket algorithm. In token bucket algorithm tokens are generated at each tick (up to certain limit). For an incoming packet to be transmitted, it must capture a token and the transmission takes place at the same rate. Hence some of the busty packets are transmitted at the same rate if tokens are available and thus introduces some amount of flexibility in the system. This also improves the performance.

4. What is choke packet? How is it used for congestion control?

Ans : Choke packet scheme is a close loop mechanism where each link is monitored to examine how much utilization is taking place. If the utilization goes beyond a

certain threshold limit, the link goes to a warning and a special packet, called **choke**
packet is sent to the source. On receiving the choke packet, the source reduced the
traffic in order to avoid congestion.

5. What is congestion? Why congestion occurs?

Ans : In a packet switching network, packets are introduced in the nodes (i.e. *offered*
load), and the nodes in-turn forward the packets (i.e. *throughput*) into the network.
When the "offered load" crosses certain limit, then there is a sharp fall in the
throughput. This phenomenon is known as **congestion**.
In every node of a packet switching network, queues (or buffers) are maintained to
receive and transmit packets (store/forward network). Due to busty nature of the
network traffic there may be situations where there is overflow of the queues. As a
result there will be re-transmission of several packets, which further increases the
network traffic. This finally leads to **congestion**.

UNIT V

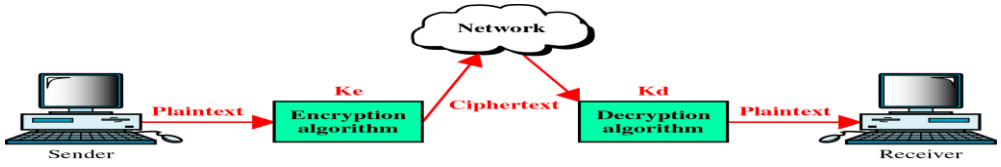
APPLICATION LAYER

Domain Name System (DNS) – E-mail – FTP – WWW – HTTP – Multimedia Network Security: Cryptography – Symmetric key and Public Key algorithms - Digital signature – Management of Public keys – Communication Security – Authentication Protocols.

5.1 Network Security

Figure 23-7

Concept of Encryption and Decryption



Specific Instructional Objectives

On completion, the students will be able to:

- State the need for secured communication
- Explain the requirements for secured communication
- Explain the following cryptographic algorithms:
- Symmetric-key Cryptography
    - Traditional ciphers
    - Monoalphabetic Substitution

- Polyalphabetic Substitution
- Transpositional Cipher
- Block ciphers Public-key Cryptography
- The RSA Algorithm

Introduction

The word **cryptography** has come from a Greek word, which means *secret writing*. In the present day context it refers to the tools and techniques used to make messages secure for communication between the participants and make messages immune to attacks by hackers. For private communication through public network, cryptography plays a very crucial role. The role of cryptography can be illustrated with the help a simple model of cryptography as shown in Fig. 8.1.1. The message to be sent through an unreliable medium is known as **plaintext**, which is encrypted before sending over the medium. The encrypted message is known as **ciphertext**, which is received at the other end of the medium and decrypted to get back the original plaintext message. In this lesson we shall discuss various cryptography algorithms, which can be divided into two broad categorize - **Symmetric key cryptography** and **Public key cryptography.**

5.2 Symmetric Key Cryptography

The cipher, an algorithm that is used for converting the plaintext to ciphertex, operates on a **key**, which is essentially a specially generated number (value). To decrypt a secret message (ciphertext) to get back the original message (plaintext), a decrypt algorithm uses a decrypt key. In symmetric key cryptography, same key is shared, i.e. the same key is used in both encryption and decryption. The algorithm used to decrypt is just the inverse of the algorithm used for encryption. For example, if addition and division is used for encryption, multiplication and subtraction are to be used for decryption. Symmetric key cryptography algorithms are simple requiring lesser execution time. As a consequence, these are commonly used for long messages. However, these algorithms suffer from the following limitations: Requirement of large number of unique keys. For example for n users the number of keys required is n(n-1)/2. Distribution of keys among the users in a secured manner is difficult

Monoalphabetic Substitution

One simple example of symmetric key cryptography is the *Monoalphabetic substitution*. In this case, the relationship between a character in the plaintext and a character in the ciphertext is always one-to-one. An example Monoalphabetic substitution is the Caesar cipher. In this approach a character in the ciphertext is substituted by another character shifted by three places, e.g. A is substituted by D. Key feature of this approach is that it is very simple but the code can be attacked very easily.

Polyalphabetic Substitution

This is an improvement over the Caesar cipher. Here the relationship between a character in the plaintext and a character in the ciphertext is always one-to-manyExample of polyalphabetic substitution is the Vigenere cipher. In this case, a particular character is substituted by different characters in the ciphertext depending on its position in the plaintext. Here the top row shows different characters in the plaintext and the characters in different bottom rows show the characters by which a particular character is to be replaced depending upon its position in different rows from row-0 to row-25. • Key feature of this approach is that it is more complex and the code is harder to attack successfully.

Transpositional Cipher

The transpositional cipher, the characters remain unchanged but their positions are changed to create the ciphertext. The characters are arranged in two-dimensional matrix and columns are interchanged according to a key is shown in the middle portion of the diagram. The key defines which columns are to be swapped. As per the key shown in the figure, character of column is to be swapped to column 3, character of column 2 is to be swapped to column 6, and so on. Decryption can be done by swapping in the reverse order using the same key. Transpositional cipher is also not a very secure approach. The attacker can find the plaintext by trial and error utilizing the idea of the frequency of occurrence of characters.

Block Ciphers

Block ciphers use a block of bits as the unit of encryption and decryption. To encrypt a 64-bit block, one has to take each of the 264 input values and map it to one of the 264 output values. The mapping should be one-to-one. Encryption and decryption operations of a block cipher are shown in Fig. 8.1.6. Some operations, such as permutation and substitution, are performed on the block of bits based on a key (a secret number) to produce another block of bits. The permutation and substitution operations. In the decryption process, operations are performed in the reverse order based on the same key to get back the original block of bits.

Transformations in Block Ciphers **Permutation:**, the permutation is performed by a permutation box at the bit-level, which keeps the number of 0s and 1s same at the input and output. Although it can be implemented either by a hardware or a software, the hardware implementation is faster. Permutation operation used in Block Ciphers **Substitution:** the substitution is implemented with the help of three building blocks – a decoder, one p-box and an encoder. For an n-bit input, the decoder produces an 2^n bit output having only one 1, which is applied to the P-box. The P-box permutes the output of the decoder and it is applied to the encoder. The encoder, in turn, produces an n-bit output. For example, if the input to the decoder is 011, the output of the decoder is 00001000. Let the permuted output is 01000000, the output of the encoder is 011. It performs the following steps: **Step-2:** Substitute each 8-bit based on **Step-3:** Permute the bits based on the key A

Figure 23-13

Permutation

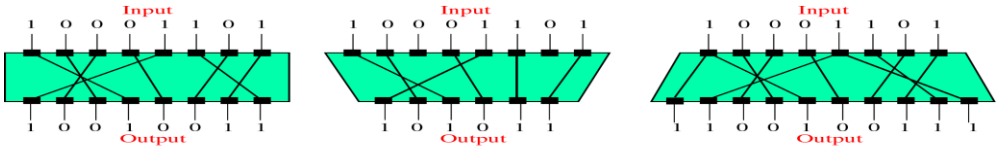


Figure 23-14

Substitution

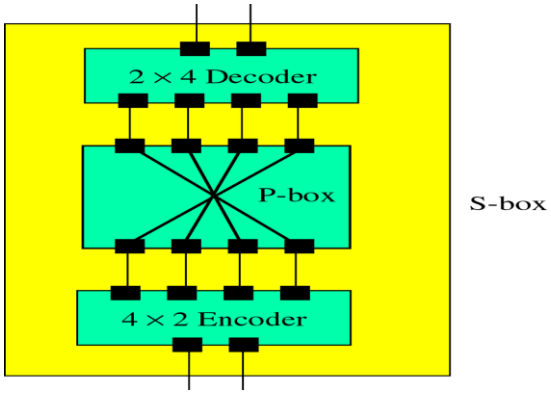


Figure 23-21

Public Key Encryption

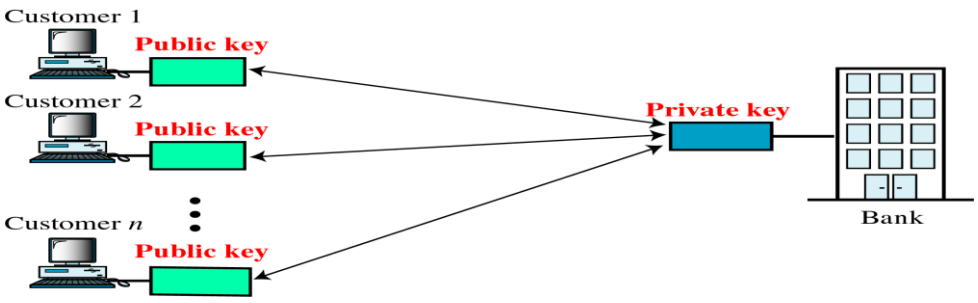


Figure 27-3

Public key encryption

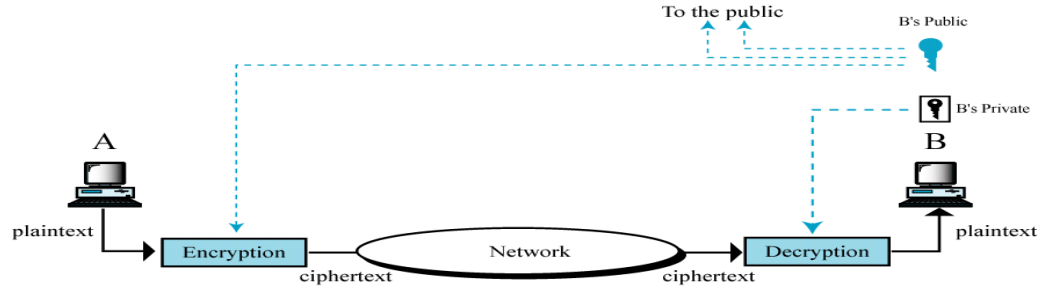
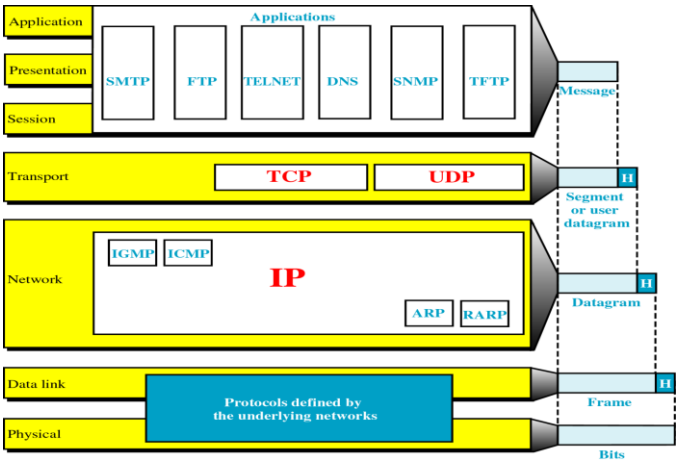


Figure 23-25

Signature Authentication

Figure 24-2

TCP/IP and the OSI Model



5.3 Encryption Standard (DES)

One example of the block cipher is the Data Encof the DES algorithm are given below: • A monoalphabetic substitution c• It has 19 distinct stages • Although the input key fonly 56 bits in length. • The decryption can be carried out in reverse order. • DES has 16 rounds, meaning the ciphertext. • As the number oexponentially. • Once the key scactual encryption or decryption is performed with the help of the main DES algorithm **Caining (CBC)** In this mode of operation, encrypted ciphernext plaintext block to be encrypted, thus making all the blocks dependent on all the previous blocksCipher Feedback Mode (CFB) encryption technique **Output Feedback Mode (OFB)** The encryption technique of Output Feedback Mode (OFB) is shown in Fig. 8.1.14. Key features of this mode are mentioned below: • OFB is also a stream cipher • Encryption is performed by XORing the message with the one-time pad • One-time pad can be generated in advance • If some bits of the ciphertext get garbled, only those bits of plaintext get garbled • The message can be of any arbitrary size • Less secure than other modes

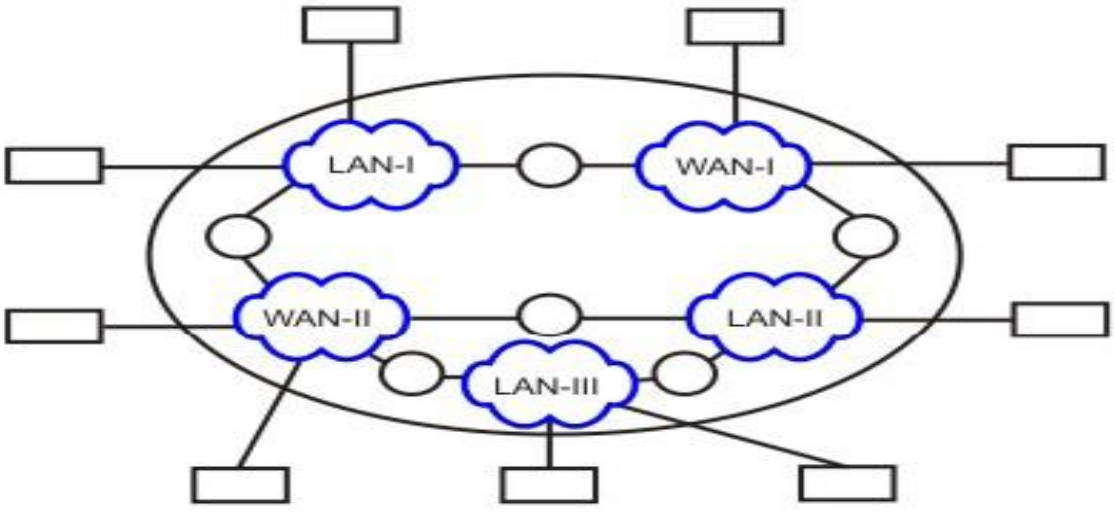
Triple DES

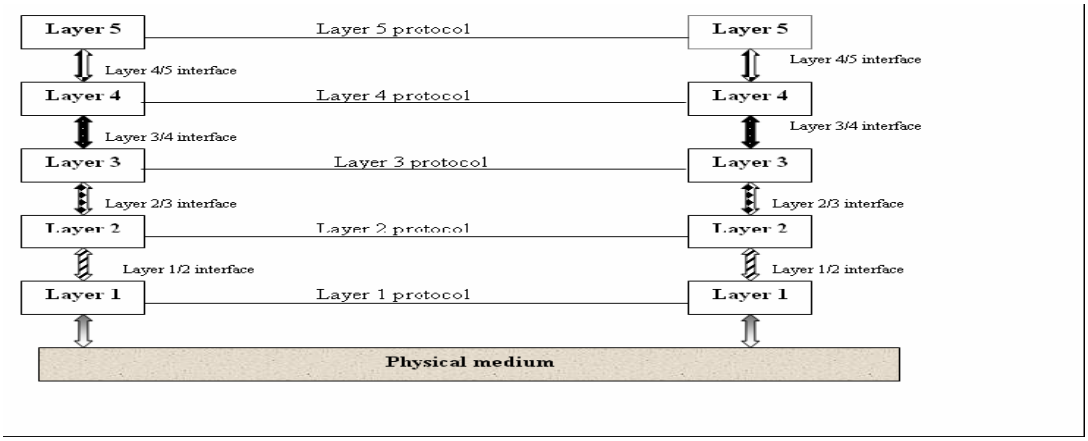
Tincreasing the key length. Its operation is explained below: • Each block of plaintext is subjected to enencryption by K1 in a sequen• CBC is used to turn the block encryption scheme into a stream encryption

5.4 Public key Cryptography

In public key cryptography, there are two keys: a private key and a public key. The public key is announced to the public, where as the private key is kept by the receiver. The sender uses the public key of the receiver for encryption and the receiver uses his private key for decryption

• Advantages: o The pair of keys can be used with any other entity The number of keys required is small o• Disadvantages: It is not efficient for long messages




Layer 5
Layer 5 protocol
Layer 5
Layer 4/5 interface
Layer 4/5 interface
Layer 4
Layer 4 protocol
Layer 4
Layer 3/4 interface
Layer 3/4 interface
Layer 3
Layer 3 protocol
Layer 3
Layer 2/3 interface
Layer 2/3 interface
Layer 2
Layer 2 protocol
Layer 2
Layer 1/2 interface
Layer 1/2 interface
Layer 1
Layer 1 protocol
Layer 1
Physical medium

Specific Instructional Objectives

On completion of this lesson, the student will be able to:

- State various services needed for secured communication
- Explain how Privacy, Authentication, Integrity and Nonrepudiation are achieved using cryptography
- State how user authentication is performed
- Explain how the PGP protocol works
- Explain how VPN works

5.5 Introduction

The basic objective is to communicate securely over an insecure medium. Any action that compromises the security of information can be considered as attack on security. Possible type of attacks mentioned below:

- **Interruption:** It is an attack on the availability of information by cutting wires, jamming wireless signals or dropping of packets by a switch.
- **Interception:** As a message is communicated through a network, eavesdroppers can listen in use it for his/her own benefit and try to tamper it.
- **Modification:** As a message is communicated through a network, eavesdroppers can intercept it and send a modified message in place of the original one.
- **Fabrication:** A message may be sent by a stranger by posing as a friend. This is also known as impersonation.

These attacks can be prevented with the help of several services implemented with the help of cryptography, as mentioned in the following section.

5.6 Security Services

Secured communication requires the following four basic services:

- **Privacy:** A person (say Sita) should be able to send a message to another person (say Ram) privately. It implies that to all others the message should be unintelligible.
- **Authentication:** After the message is received by Ram, he should be sure that the message has been sent by nobody else but by Sita.
- **Integrity:** Ram should be sure that the message has not been tampered on transit.
- **Nonrepudiation:** Ram should be able to prove at a later stage that the message was indeed received from Sita.

5.7 Privacy

Privacy can be achieved using symmetric key cryptography. In this case, the key is shared between the sender (Sita) and the receiver (Ram) as shown in Fig. 8.2.1. Privacy can also be achieved by using public-key cryptography as shown in Fig. 8.2.2. However, in this case the owner should be verified.

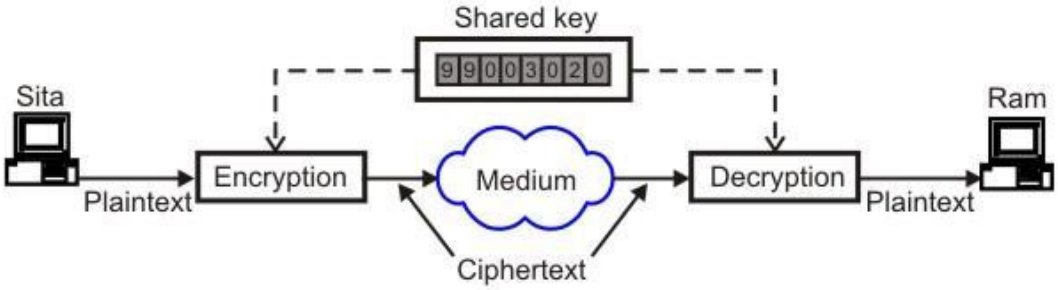


Figure 8.2.1 Privacy using private-key cryptography

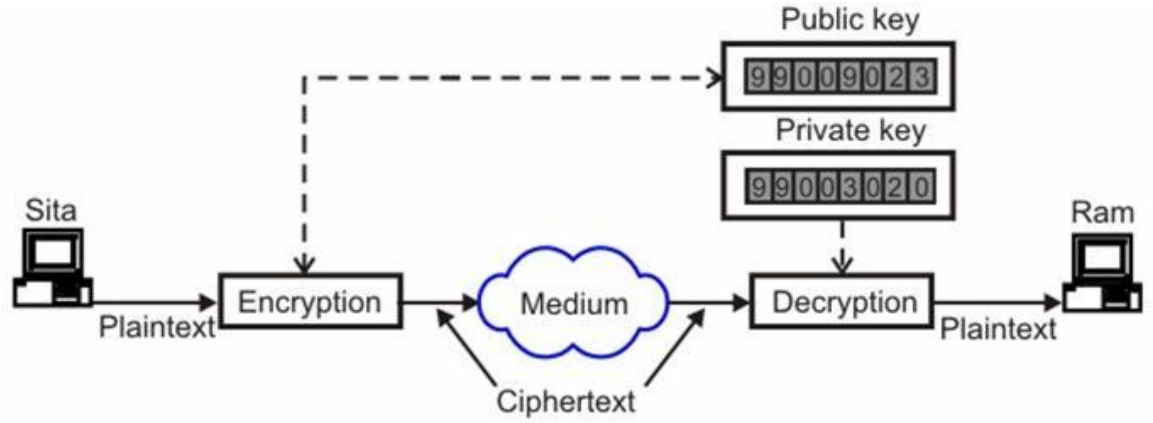


Figure 8.2.2 Privacy using public-key cryptography.

5.8 Authentication, Integrity and Nonrepudiation using Digital Signature

By message authentication we mean that the receiver should be sure about sender's identity. One approach to provide authentication is with the help of digital signature. The idea is similar to signing a document. Digital Signature provides the remaining three security services; Authentication, Integrity and Nonrepudiation.

Digital Signature

There are two alternatives for Digital Signature:

- Signing the entire document
- Signing the digest

In the first case the entire document is encrypted using private key of the sender and at the receiving end it is decrypted using the public key of the sender as shown in Fig. 8.2.3. For a large message this approach is very inefficient. In the second case a miniature version of the message, known as *digest*, is encrypted using the private key of the sender and then the signed digest along with the message is sent to the receiver as shown in Fig. 8.2.4. The receiver decrypts the signed digest using the public key of the sender and the digest created using the received message is compared with the decrypted digest as shown in Fig. 8.2.5. If the two are identical, it is assumed that the sender is authenticated.This is somewhat similar to error detection using parity bit.

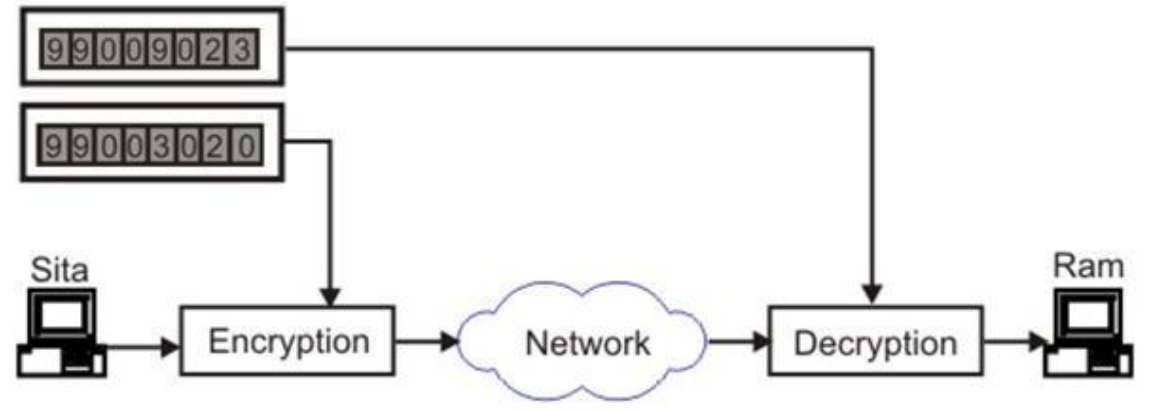


Figure 8.2.3 Authentication by signing the whole document.

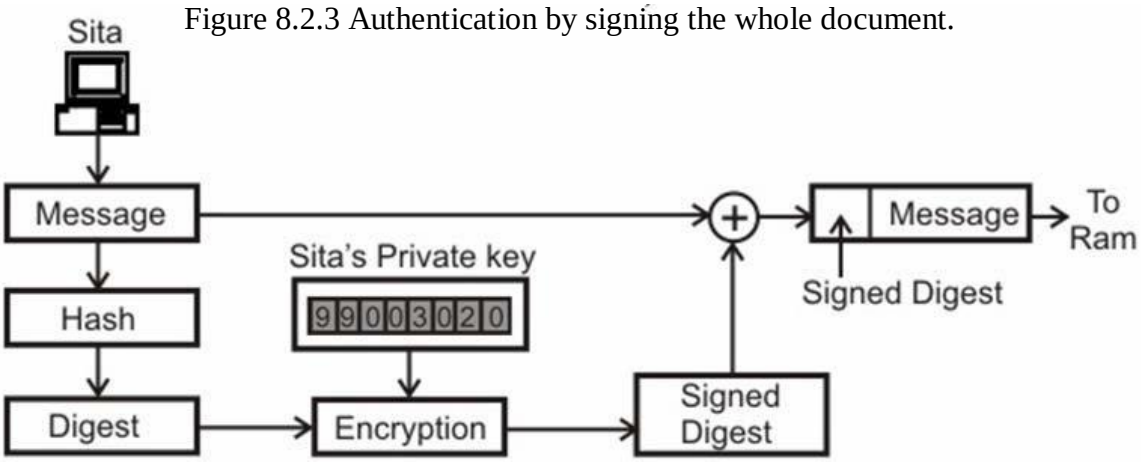


Figure 8.2.4 Sender site for authentication by signed digest

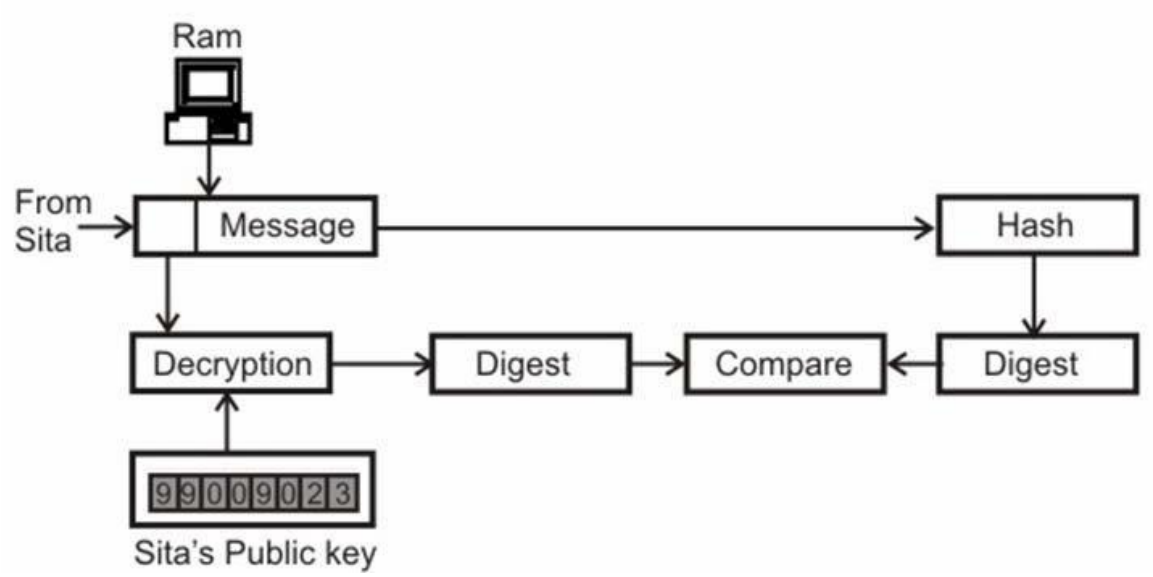


Figure 8.2.5 Receiver site for authentication by signed digest

Some key features of this approach are mentioned below:

- Digital signature does not provide privacy
- Hash function is used to create a message digest
- It creates a fixed-length digest from a variable-length message
- Most common Hash functions:
  - MD5 (Message Digest 5): 120-bit
  - SHA-1 (Secure Hash algorithm 1): 160-bit
- Important properties:
- One-to-One
- One-way

5.9 User Authentication using symmetric key cryptography

User authentication is different from message authentication. In case of message authentication, the identity of the sender is verified for each and every message. On the other hand, in user authentication, the user authentication is performed once for the duration of system access. In the first approach, the sender (Sita) sends her identity and password in an encrypted message using the symmetric-key K_{SR} and then sends the message as shown in Fig. 8.2.6. However, an intruder (say Ravana) can cause damage without accessing it. He can also intercept both the authentication message and the data message, store them and then resends them, which is known as *replay attack*.

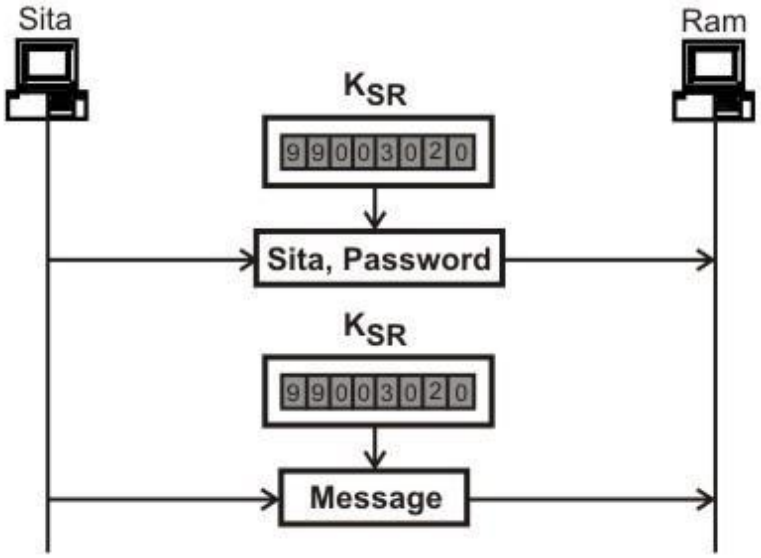


Figure 8.2.6 User authentication using symmetric key cryptography

Using nonce, a large random number used only once

To prevent the replay attack, the receiver (Ram) sends *nonce*, a large random number that is used only once to the sender (Sita) to challenge Sita. In response Sita sends an encrypted version of the random number using the symmetric key. The procedure is shown in Fig. 8.2.7.

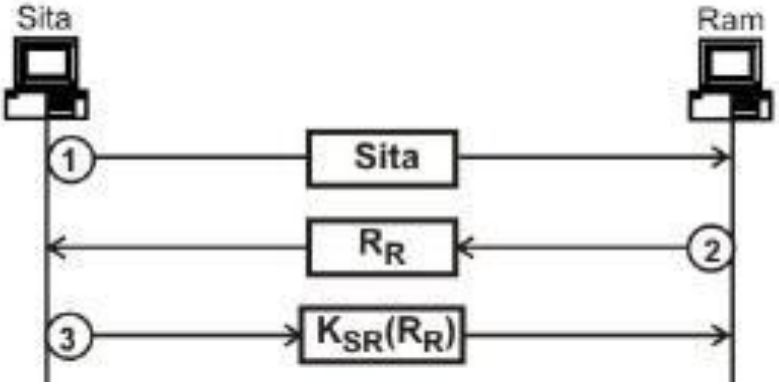


Figure 8.2.7 User authentication using a nonce

Bidirectional Authentication

In the bidirectional authentication approach, Ram sends *nonce* to challenge Sita and Sita in turn sends nonce to challenge Ram as shown in Fig. 8.2.8. This protocol uses extra messages for user authentication. Protocol with lesser number of messages is possible as shown in Fig. 8.2.9.

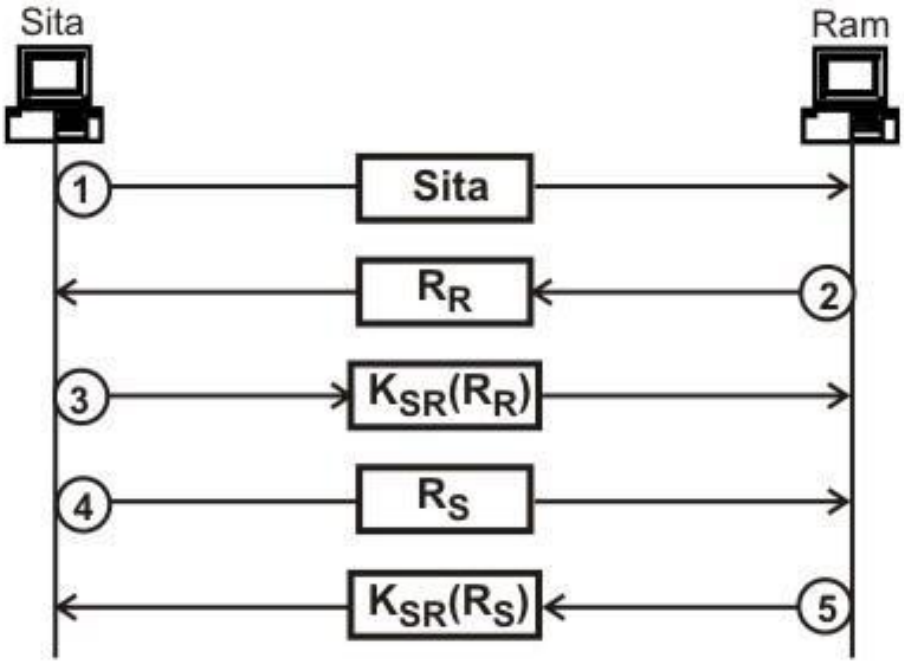


Figure 8.2.8 Bidirectional authentication using a nonce

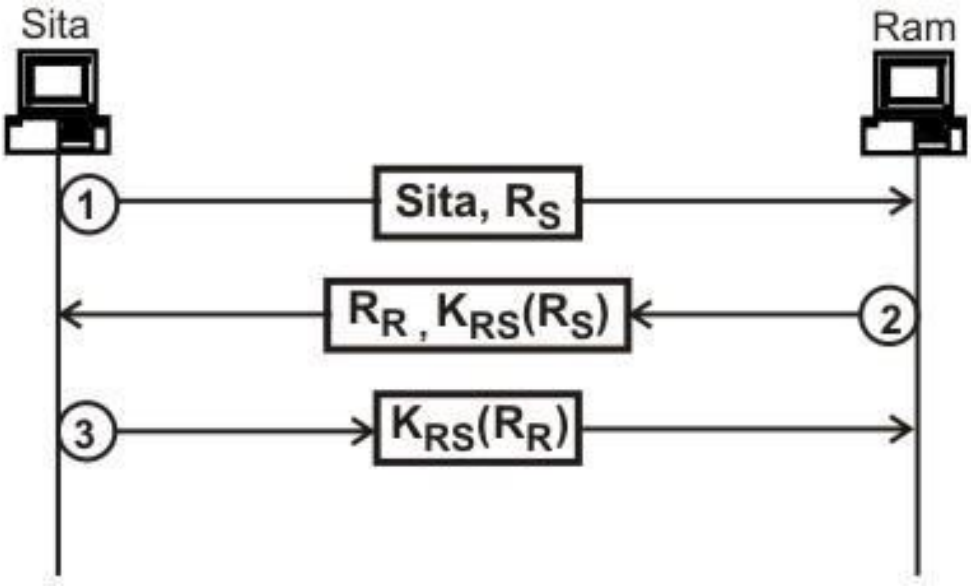
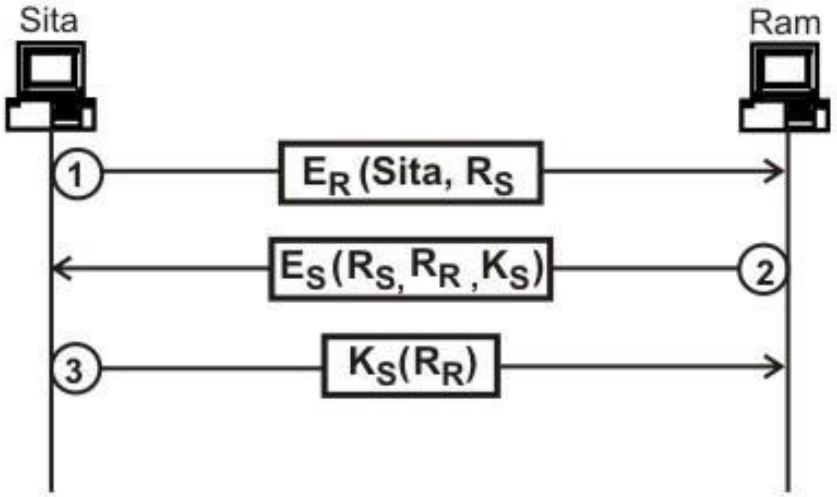


Figure 8.2.9 Bidirectional authentication using lesser number of messages

5.10 User Authentication using Public Key Cryptography

Public key cryptography can also be used to authenticate a user. The procedure is shown in Fig. 8.2.10.



E_R= Public key of Ram, E_S= Public key of Sita
R_S= nonce by Sita, R_R= nonce by Ram
K_s= Session key sent by Ram

Figure 8.2.10 User authentication using public key cryptography

5.11 Key Management

Although symmetric-key and public-key cryptography can be used for privacy and user authentication, question arises about the techniques used for the distribution of keys. Particularly, symmetric-key distribution involves the following three problems:

- For n people to communicate with each other requires n(n-1)/2 keys. The problem is aggravated as n becomes very large.
- Each person needs to remember (n-1) keys to communicate with the the remaining (n-1) persons.
- How the two parties will acquire the shared key in a secured manner?

In view of the above problems, the concept of *session key* has emerged. A session key is created for each session and destroyed when the session is over. The **Diffie-Hellman** protocol is one of the most popular approach for providing one-time session key for both the parties.

Diffie-Hellman Protocol

Key features of the Diffie-Hellman protocol are mentioned below and the procedure is given in Fig. 8.2.11.

- Used to establish a shared secret key
- Prerequisite: N is a large prime number such that (N-1)/2 is also a prime number. G is also a prime number. Both N and G are known to Ram and Sita..
- Sita chooses a large random number x and calculates R1 = G^x mod N and sends it to Ram
- Ram chooses another large random number y and calculates R2 = G^y mod N and sends it to Sita
- Ram calculates K = $(R1)^y$ mod N
- Sita calculates K = $(R2)^x$ mod N

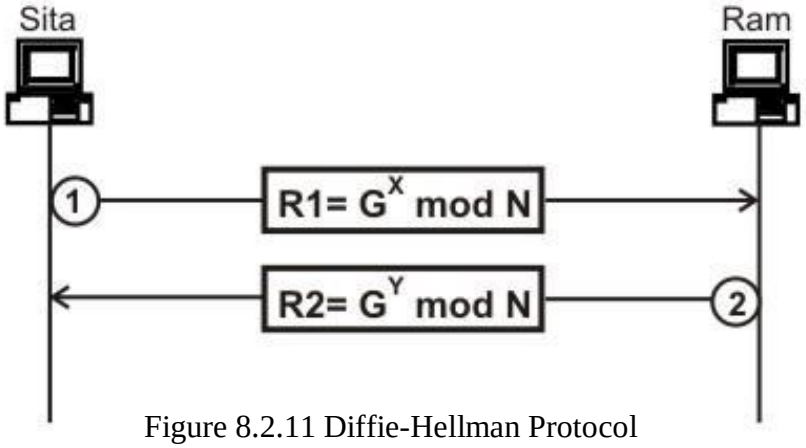


Figure 8.2.11 Diffie-Hellman Protocol

Key Management using KDC

It may be noted that both R1 and R2 are sent as plaintext, which may be intercepted by an intruder. This is a serious flaw of the Diffie-Hellman Protocol. Another approach is to use a trusted third party to assign a symmetric key to both the parties. This is the basic idea behind the use of *key distribution center (KDC).*

Key Management using Kerberos

Another popular authentication protocol known as *Kerberos* It uses an authentication server (AS), which performs the role of KDC and a ticket-granting server (TGS), which provides the session key (K_{AB}) between the sender and receiver parties. Apart from these servers, there is the real data server say Ram that provides services to the user Sita. The operation of Kerberos is depicted with the help of Fig. 8.2.12. The client process (Sita) can get a service from a process running in the real server Ram after six steps as shown in the figure. The steps are as follows:
Step 1. Sita uses her registered identity to send her message in plaintext.
Step 2. The AS server sends a message encrypted with Sita's symmetric key K_s. The message contains a session key K_{se}, which is used by Sita to contact the TGS and a ticket for TGS that is encrypted with the TGS symmetric key K_{TG}.

Step3. Sita sends three items to the TGS; the ticket received from the AS, the name of the real server, and a timestamp encrypted by K_{se}. The timestamp prevents replay by Ram.
Step 4. The TGS sends two tickets to Sita. The ticket for Sita encrypted with Kse and the ticket for Ram encrypted with Ram's key. Each of the tickets contains the session key K_{SR} between Sita and Ram.
Step 5. Sita sends Ram's ticket encrypted by K_{SR}.
Step 6. Ram sends a message to Sita by adding 1 to the timestamp confirming the receipt of the message using K_{SR} as the key for encryption.

Following this Sita can request and get services from Ram using K_{SR} as the shared key.

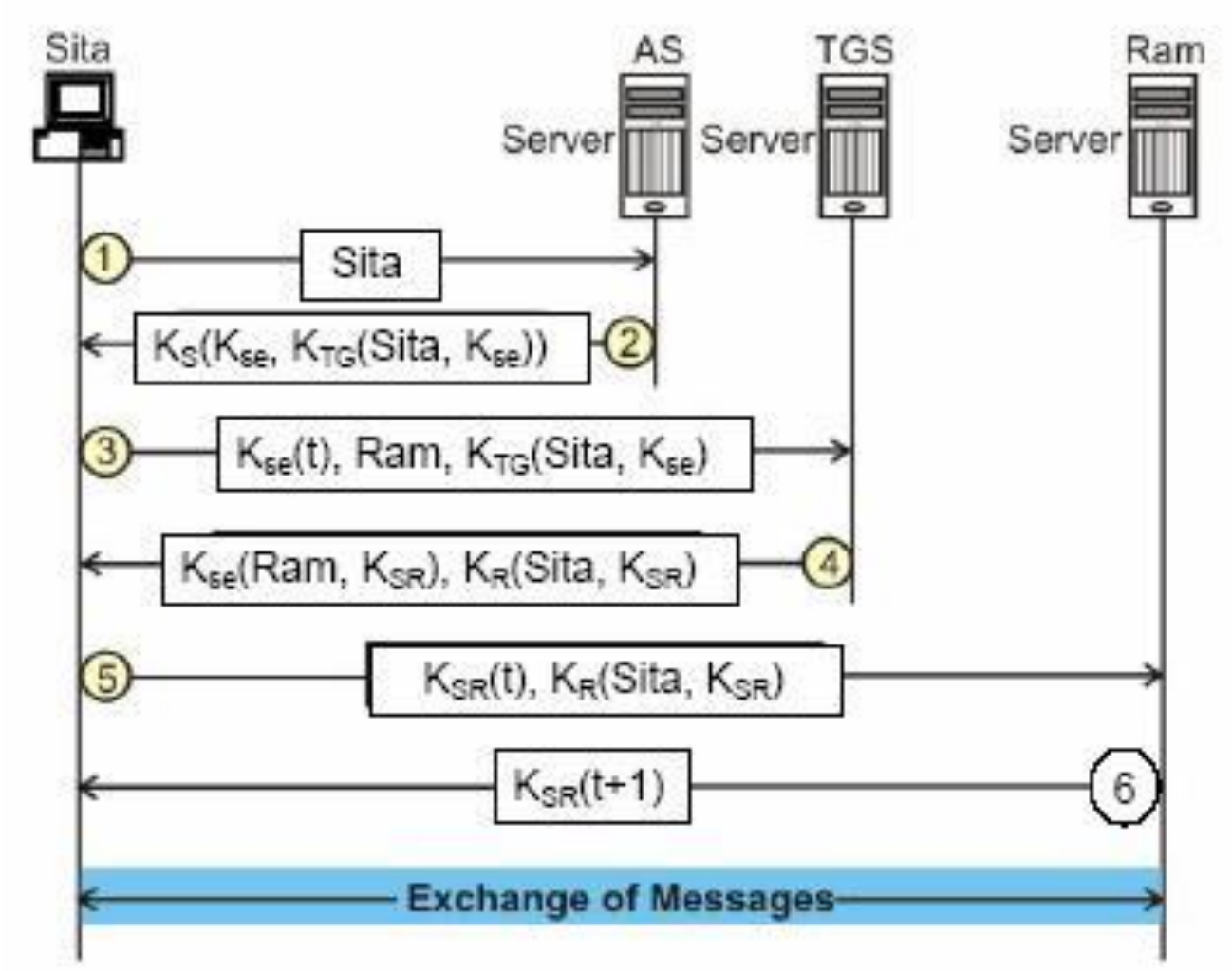


Figure 8.2.12 The Kerberos Protocol

5.12 Application Layer Security

Based on the encryption techniques we have discussed so far, security measures can be applied to different layers such as network, transport or application layers. However, mplementation of security features in the application layer is far simpler and feasible ompared to implementing at the other two lower layers. In this subsection, a protocol known as *Pretty Good Privacy (PGP)*, invented by Phil Zinmermann, that is used in the application layer to provide all the four aspects of security for sending an email is briefly discussed. PGP uses a combination of private-key and public key for privacy. For integrity, authentication and nonrepudiation, it uses a combination of hashing to create digital signature and public-key encryption as shown in Fig. 8.2.13.

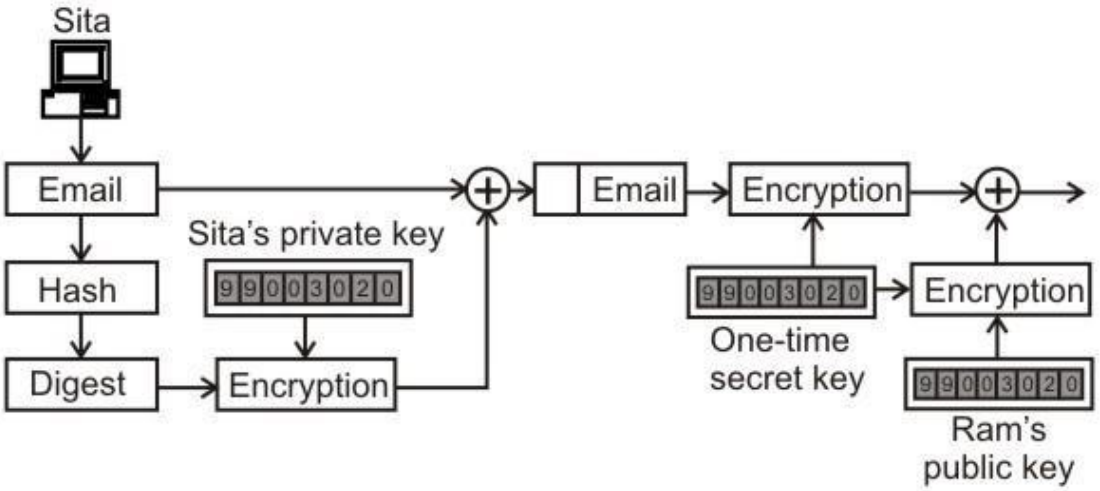


Figure 8.2.13 (a) Sender site of the PGP

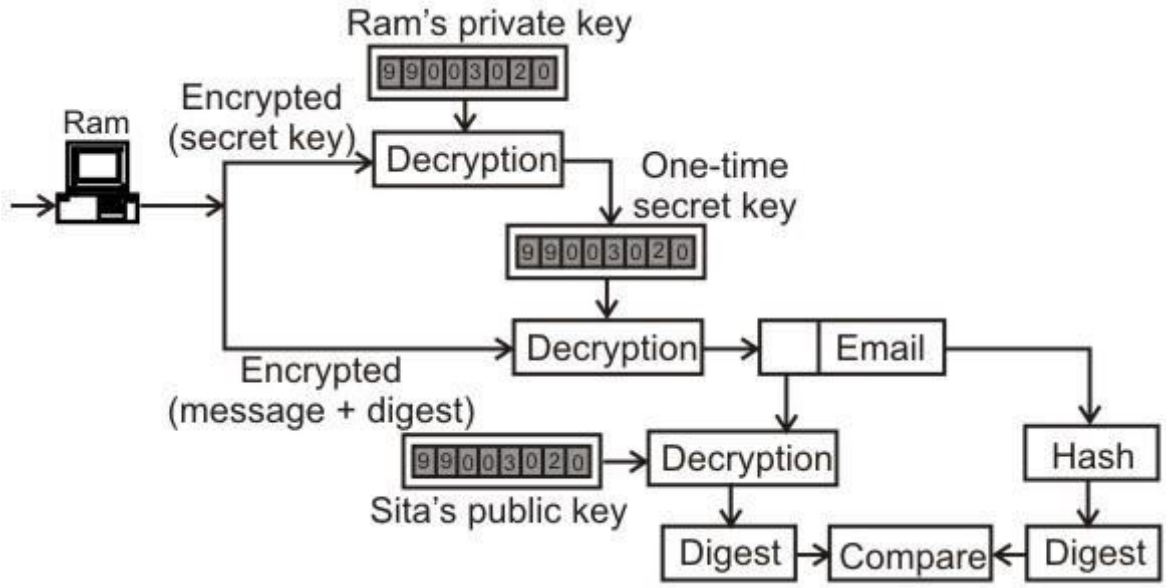


Figure 8.2.13 (b) Receiver site of the PGP

5.13 Virtual Private Network (VPN)

With the availability of huge infrastructure of public networks, the *Virtual Private Network (VPN)* technology is gaining popularity among enterprises having offices distributed throughout the country. Before we discuss about the VPN technology, let us first discuss about two related terms: *intranet* and *extranet*.

Intranet is a private network (typically a LAN) that uses the internet model for exchange of information. A private network has the following features:

- It has limited applicability because access is limited to the users inside the network
- Isolated network ensures privacy
- Can use private IP addresses within the private network

Extranet is same as the intranet with the exception that some resources can be allowed to access by some specific groups under the control of network administrator.

Privacy can be achieved by using one of the three models: Private networks, Hybrid Networks and Virtual Private Networks.

Private networks: A small organization with a single site can have a single LAN whereas an organization with several sites geographically distributed can have several LANs connected by leased lines and routers as shown in Fig. 8.2.14. In this scenario, people inside the organization can communicate with each other securely through a private internet, which is totally isolated from the global internet.

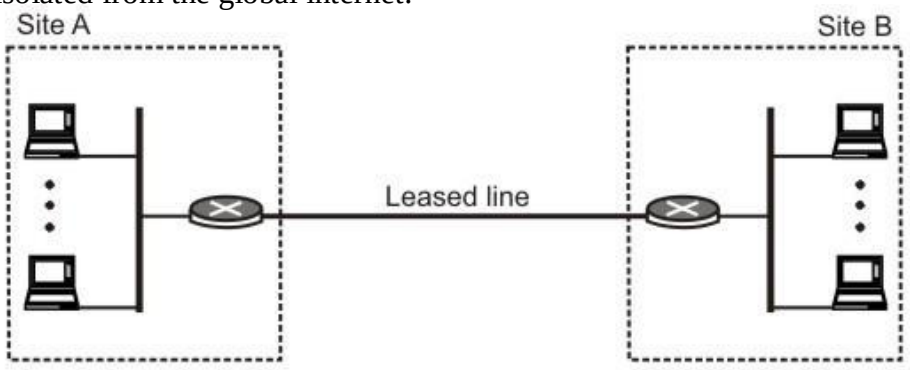


Figure 8.2.14 Private network with two LAN sites

Hybrid Networks: Many organizations want privacy for inter-organization level data exchange, at same time they want to communicate with others through the global internet. One solution to achieve this is to implement a hybrid network as shown in Fig. 8.2.15. In this case, both private and hybrid networks have high cost of implementation, particularly private WANs are expensive to implement.

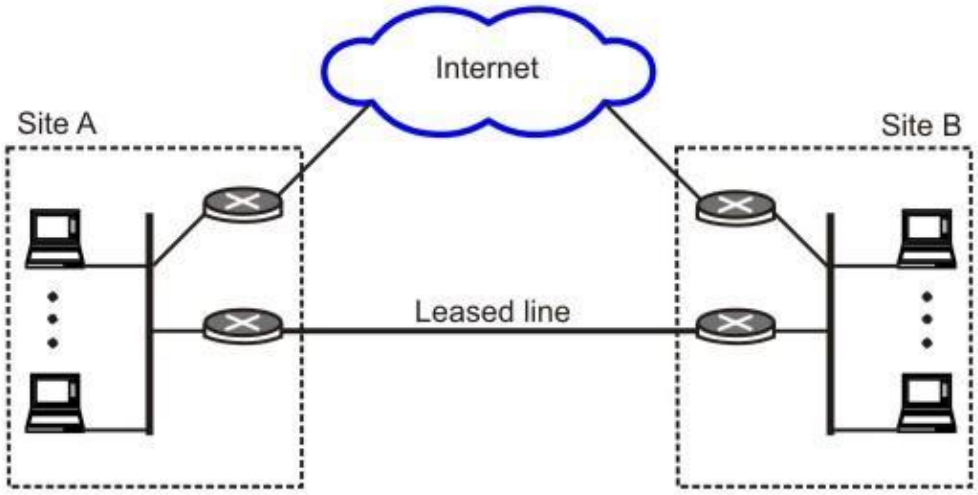


Figure 8.2.15 Hybrid network with two LAN sites

Virtual Private Networks (VPN):

VPN technology allows both private communication and public communications through the global internet as shown in Fig. 8.2.16. VPN uses IPSec in the tunnel mode to provide authentication, integrity and privacy. In the IPSec tunnel mode the datagram to be sent is encapsulated in another datagram as payload. It requires two sets of addressing as shown in Fig. 8.2.17.

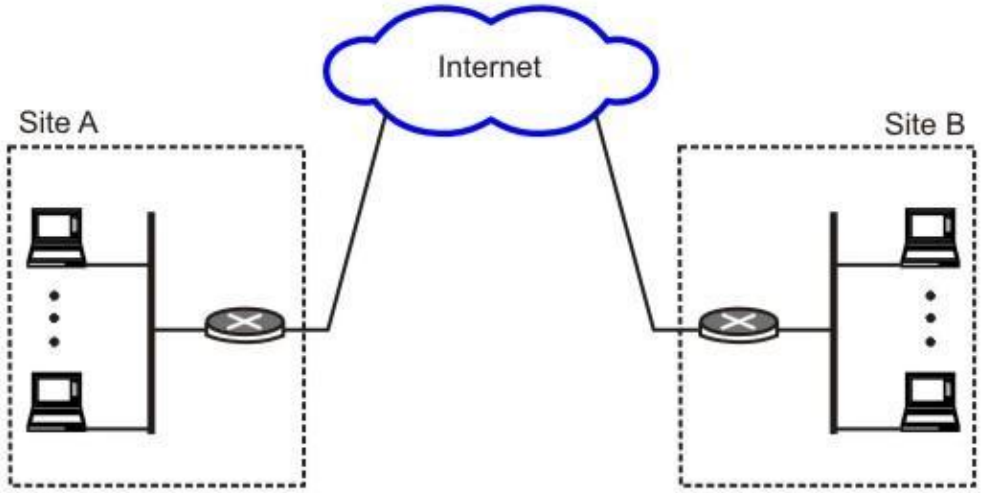
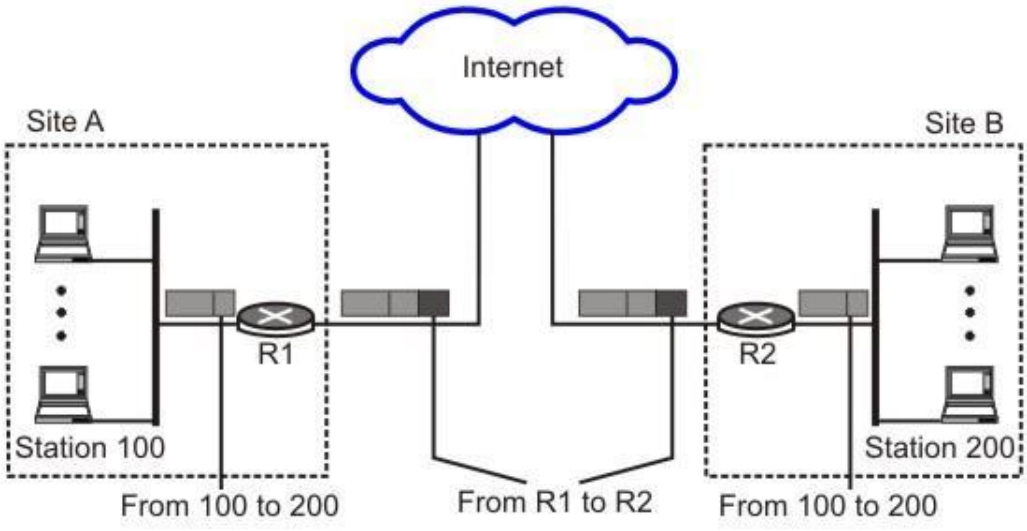


Figure 8.2.16 VPN linking two LANs



New IP Header **IPSec Header** **IP Header** **Rest of the packet**

TWO MARKS QUESTIONS

1.What are the four services required for secured communication?

Ans: The four services required for secured communication are: privacy, integrity, authentication and nonrepudiation.

2. What is nonce?

Ans: The nonce is a large random number that is used only once for the purpose of user
authentication.

3. Explain the operation of the Diffie-Hellman protocol with an example.

Ans: Although the algorithm works on l;arge numbers, it is illustrated with smaller numbers in this example.
Let N = 23 and G = 7.
Sita chooses x = 5 and calculates R1 = 7^5Mod23 = 17
Sita sends 17 to Ram.
Ram chooses y = 3 and calculates R2 = 7^3Mod23 = 21
Ram sends 21 to Sita
Ram calculates K = 17^3Mod 23 = 14
Sita Calculates K = 21^5Mod23 = 14

4. Explain the function of Kerberos.

Ans: Kerberos is a popular technique for key distribution. The kerberos is an authentication protocol and at the same time acts as a Key Distribution Center. It requires an authentication server and a ticket-granting server in addition to the real data server.

5. What is VPN?

Ans: VPN allows private communication through public internet. It is essentially a logical (virtual) network within a conventional network. It makes use of cryptography (IPSec in tunnel mode) to perform private communication through insecure and public internet.